PREDICTION MODELING FOR GRADUATE ATHLETIC

TRAINING EDUCATION PROGRAMS

By

Scott L. Bruce, MS, ATC

Elizabeth Crawford
Assistant Professor
(Chair)

Martina Harris
Assistant Professor
(Committee Member)

Gary B. Wilkerson
Professor
(Committee Member)

David Rausch
Associate Professor
(Committee Member)

R. Barry Dale
Associate Professor
(Committee Member)

PREDICTION MODELING FOR GRADUATE ATHLETIC

TRAINING EDUCATION PROGRAMS

By

Scott L. Bruce, MS, ATC

A Dissertation Submitted to the Faculty of the
University of Tennessee at Chattanooga
in Partial Fulfillment of the
Requirements of the Degree of
Doctor of Education

The University of Tennessee at Chattanooga
Chattanooga, Tennessee

May 2014

ABSTRACT

The purposes of this study were:

- to develop a prediction model to identify factors associated with eligibility and first-attempt success on the Board of Certification (BOC) examination for students enrolled in a professional (entry-level) graduate athletic training program (GATP)
- to identify applicant characteristics that are most likely to predict both academic success in the GATP and success on the BOC exam. A cohort of 119 students was used for both purposes. Multiple analyses yielded three-factor and two-factor models for prediction of passing the BOC exam.

The three-factor model demonstrates that a student with ≥ 2 predictors had an odds ratio (OR) of 6.31 and a relative frequency of success (RFS) 1.66 for passing the BOC exam and correctly predicted 87.7% of first-attempt success on the BOC exam. The two-factor model demonstrates a student with ≥1 predictor had an OR of 10.69 and an RFS 2.05 for passing the BOC exam and correctly predicted 89.2% of first-attempt success on the BOC exam.

Multiple analyses yielded two three-factor models for prediction of success in the GATP. The initial three-factor model demonstrates that a student with ≥2 predictors had an OR of 17.94 and a RFS of 2.13 for students being successful in the GATP, and correctly predicted 90.5% of GATP success. The alternative three-factor model found a student with ≥2 predictors had an OR

of 20.94 and an RFS 1.98 for students being successful in the GATP, and correctly predicted 93.9% of GATP success.

Within the past year, changes in athletic training education have been implemented and more are expected in the future, specifically whether or not a graduate professional (entry-level) athletic training degree will be required to sit for the BOC exam. Since there is a greater emphasis on first-time BOC exam pass rates, and more programs convert to graduate level curricula, the results of this study may assist GATPs to identify students who are likely to be successful in the graduate program and to pass the BOC exam on the first-attempt.

DEDICATION

I dedicate this dissertation to my wonderful wife, Jana Bruce, and to my two outstanding children, Patrick and Allison. Without their love and support, completion of this dissertation and degree would not have been possible. Their sacrifice and understanding is so very much appreciated. The love and respect I have for my wife is immeasurable. The character my children have developed, and the people they have become during my doctoral process is something which will always bring me a great deal of pride and happiness. Thank you all! You mean more to me than you will ever know or I can ever express.

I also dedicate this dissertation to my parents Bob and Shirley Bruce. Neither one of them is with me today, but I know they are both looking down upon me with pride. Thank you for the way you raised me. You did alright!

ACKNOWLEDGEMENTS

Firstly, I would like to acknowledge and express my appreciation to my dissertation committee chair, Dr. Elizabeth Crawford. She helped me to learn a great deal about the doctoral process, about myself, and about learning and leadership. Her helpful advice while allowing me to make mistakes and learn from them is greatly appreciated. It has truly been a privilege and honor to work with someone who is as dedicated to the learner as she is to the concepts of learning and leadership.

I would also like to express my sincere gratitude and appreciate to my methodologist and statistician Dr. Gary Wilkerson. He has been with me for most of this entire doctoral journey. He has taught me more than I could have ever learned on my own. Furthermore, he has challenged me to become a better person, a better student, a better educator, and a better clinician time and time again. The patience he has shown with me along the way, (when I did not always deserve them) is unquestionably something I will not forget, but will always appreciate.

Thank you also to my other committee members. I thank Dr. David Rausch for his sage advice, leadership, and guidance throughout the doctoral program. It is hard to believe how far the Learning and Leadership Program has come in such a short time under his direction. I have benefitted greatly from his mentorship.

To Dr. Martina Harris and Dr. Barry Dale, thank you for all of your hard work and assistance throughout my pursuit of my doctoral degree. Thank you for your willingness to be a part of my committee and helping to finish the dissertation process.

In addition I would like to thank several others for their help and support. To Dr. Marisa Colston, thank you for your support and assistance. I know there have been trying times as I struggled to complete my degree, but my goal has always been to help make the UTC GATP the very best program in the country. We are the number one GATP because of your leadership.

To Dr. Michael Biderman, whose formula not only helped to make this project possible, but also made for a successful dissertation. Thank you for your input on the results section of this study. I really appreciate your advice.

I offer my appreciation to my good friend Dr. Craig Denegar. Years ago he told me that if I ever got a doctorate to be someone who used the degree and didn't just have the degree. Later on, he showed confidence in my abilities to complete doctoral work with several others did not believe in me. There have been many times when I did not believe I would be where I am now, but as I have gone through the ups and downs of my doctoral journey, your words and confidence in me were always heard in the back of my mind. I hope to be able to use this degree to advance the profession of athletic training and can become half the professional you have been and continue to be.

Thank you also to Ms. Allie Jenkins. Your assistance in teaching me the ins and outs of the University's software system to aid in my data collection was a tremendous help and assistance. Without your help I don't know what I would have done. I would probably still be searching for the data you helped me collect quickly and efficiently.

Thank you to Dr. Gary Liguori for helping me to prioritize my responsibilities for this academic year. I appreciate the support and advice.

To another strong supporter and friend, Dr. George Mihalik, Head Football Coach at Slippery Rock University; I cannot thank you enough. Your continued encouragement and positive comments were always appreciated. We won a lot of games together and I regard this accomplishment just like any of those championships. I am glad you were on my team!

To Cohort 7 – we came into the Learning and Leadership Program together and you have taught me far more than I could have ever taught you. The respect and admiration I have for each of you is immense. We might not be leaving together, but know as each of you completes his or her journey, a little bit of the cohort is going with you. Stick to it and finish. If I can do it, then you certainly can do it.

Thank you to Ms. Marti Wayland. You are the reason I passed and understood my statistics class. Thank you for tutoring and encouraging me!

I acknowledge Mr. Vince Sinovic who got me involved in athletic training in 1976. You have been like a second father to me. We shared a lot of good times together. I only hope I have made you proud.

To my other family and friends I wish to express my heartfelt appreciation for keeping me humble and honest and the answer is no, most of you will not have to refer to me as doctor.

Finally, and most importantly, I thank God and Jesus Christ my Lord and Savior. Jeremiah 29:11 has never meant so much as it did during this process. Thank you!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACT, American College Test

ATC, Certified Athletic Trainer

AUC, Area Under the Curve

BD, Breslow-Day

BOC, Board of Certification

CI, Confidence Interval

Exp(B), Exponent B or adjusted odds ratio

GATP, Graduate Athletic Training Program

gGPA, Graduate Grade Point Average

GPA, Grade Point Average

GRE, Graduate Record Examination

GREq, Quantitative section of the Graduate Record Examination

GREv, Verbal section of the Graduate Record Examination

GREwr, Analytical Writing section of the Graduate Record Examination

MCAT, Medical College Admission Tests

MH, Mantel- Haenszel

NPTE, National Physical Therapy Examination

OR, Odds Ratio

$OR_{MH,}$ Odds Ratio for the Mantel-Haenszel estimator equation

OT, Occupational Therapist

PCTL, Percentile

PR, Percentile Rank

GREq PR, Percentile Rank of the Quantitative section of the Graduate Record Examination

GREv PR, Percentile Rank of the Verbal section of the Graduate Record Examination

GREwr PR, Percentile Rank of the Analytical Writing section of the Graduate Record Examination

PT, Physical Therapist

ROC, Receiver Operator Characteristic curve

RFS, Relative Frequency for Success

SAT, Standardized Achievement Test

Sn, Sensitivity

Sp, Specificity

uGPA, Undergraduate Grade Point Average

LIST OF SYMBOLS

$\geq$, greater than or equal to

$\leq$, less than or equal to

$<$, less than

$=$, equals to

$\sum$, sum of

CHAPTER I

INTRODUCTION

A common goal of professional education programs is to recruit the best students. Selection of students from a pool of candidates can be a difficult process, especially if the number of qualified candidates exceeds the number of available positions. How these decisions are made and who to include or exclude can be difficult, and they can be even more difficult to defend once they are made. The more objective the selection process the easier it can be to identify qualified candidates and to defend against any legal actions or other potential problems related to candidates not accepted into the program.

Multiple health education program administrators have examined potential predictors for assisting in their decisions to admit or reject students. A literature search on programs from clinical psychology, nursing, occupational therapy, physician assistant, physical therapy, and medical school found all have attempted to refine their selection processes (Balogun, Karacoloff, & Farina, 1986; Ferguson, James, & Madeley, 2002; Hansen & Pozehl, 1995; Hayes, Fiebert, Carroll, & Magill, 1997; Kirchner, Holm, Ekes, & Williams, 1994; Levine, Knecht, & Eisen, 1986; McGinnis, 1984; Meleca, 1995; Morris & Farmer, 1999; Munro, 1985; Payton, 1997; Salvatori, 2001; Vendrely, 2007; Willingham, 1972; Wilson, 1999; Zipp, Ruscingno, & Olson, 2010). Several different approaches have been used in an effort of trying to isolate which variable or group of variables are best at predicting those candidates that should be selected for their programs. Predictor variables such as the Graduate Record Examination (GRE), undergraduate grade point average (uGPA), Medical College Admission Tests (MCAT), past clinical experience, age, race, gender, and ethnicity have all been employed (Balogun et al., 1986; Ferguson et al., 2002; Hansen & Pozehl, 1995; Hayes et al., 1997; Kirchner & Holm, 1997; McGinnis, 1984; Meleca, 1995; Munro, 1985; Salvatori, 2001; Willingham, 1972; Zipp et al., 2010). Several other more subjective variables were utilized to measure successful candidates and have included written essays, interviews, subjective inventories, references, and personal characteristics. The outcome variables that have been used include admission into the program, graduate grade point average, academic performance, clinical rotation success, and graduation from the program (Balogun et al., 1986, p. 50; Bretz, 1989; Burton & Wang, 2005; Day, 1986; DeAngelis, 2003; Feldman, 2007; Ferguson et al., 2002; Hayes et al., 1997; Keskula, Sammarone, & Perrin, 1995; Kirchner & Holm, 1997; Kirchner et al., 1994; McGinnis, 1984; Meleca, 1995; Mitchell, 1990; Munro, 1985; Payton, 1997; Platt, Sammarone-Turocy, & McGlumphy, 2001; Sime, Corcoran, & Libera, 1983; Utzman, Riddle, & Jewell, 2007a; Willingham, 1972; Zipp et al., 2010).

No studies to date have examined admittance decisions for graduate professional (entry-level) athletic training education programs. Keskula, Sammarone & Perrin (1995) conducted a study to examine prediction variables for post-professional NATA-approved Graduate Athletic Training Program using stepwise multiple regression analysis to determine that uGPA was the only significant predictor of graduate school GPA (Keskula et al., 1995).

Medical programs such as anesthesiology, athletic training, medicine and physical therapy have all attempted to predict success on their respective licensing and board examinations (Armstrong, Dahl, & Haffner, 1998; Kosmahl, 2005; Kuncel, Wee, Serafin, & Hezlett, 2010; McClintock & Gravlee, 2010; Utzman, Riddle, & Jewell, 2007b; Zaglaniczny, 1992). There have been nine studies attempting to predict success on the National Athletic Trainers' Association's Board of Certification (BOC) examination (Draper, 1989; Erickson & Martin, 2000; Harrelson, Gallaspy, Knight, & Leaver-Dunn, 1997; Hickman, 2010; Middlemas, Manning, Gazzillo, & Young, 2001; Pickard, 2003; Starkey & Henderson, 1995; Turocy, Comfort, Perrin, & Gieck, 2000; Williams & Hadfield, 2003). All nine studies examined undergraduate students in their attempts to predict success on the BOC exam; none of these studies were successful in predicting candidates' success on the BOC exam.

Statement of the Problem

There are two interrelated purposes for this study, both of which pertained to the process of admitting students to a graduate professional program. The first component of this study involves the development of a prediction model to identify factors associated with eligibility and first-attempt success on the Board of Certification (BOC) examination for students who have completed a professional (entry-level) graduate athletic training program (GATP). The second component will utilize the results of the first analysis to identify program applicant characteristics that are most likely to predict both academic success within the graduate professional program and subsequent success on the BOC exam.

Significance of the Problem

This study may serve to identify methods to aid in the selection of potential students for athletic training educational programs, thus, improving the success rate of first-time pass rate on the BOC exam of student coming from a GATP. Additionally, this study may assist athletic training education program directors to improve the quality of the educational experience for the students and permit the program director to provide sufficient advice on when the student may find the greatest likelihood for success on the BOC exam.

Hypothesis

The hypotheses for this study are first to develop a prediction model to identify factors associated with eligibility and first-attempt success on the Board of Certification (BOC) examination for students who have completed a professional (entry-level) GATP. The second component will utilize the results of the first analysis to identify program applicant characteristics that are most likely to predict both academic success within the graduate professional program and subsequent success on the BOC exam.

The first null hypothesis for this study is that a prediction model cannot be created to identify factors associated with eligibility and first-attempt success on the Board of Certification (BOC) examination for students who have completed a professional (entry-level) GATP. The second null hypothesis for this study is that the results of the first analysis cannot be used to identify program applicant characteristics that are most likely to predict both academic success within the graduate professional program and subsequent success on the BOC exam.

Outcome Variables

The outcome variables for this study are the dichotomy between passing versus not passing of the BOC exam on the first attempt taking the exam by students selected for a GATP and program academic success as measured by gGPA at the end of the first year.

Predictor Variables

The predictor variables for this study include the following:

The predictor variables for this study include the following:

- undergraduate Grade Point Average (uGPA)
- percentile rank of the GRE verbal score (GREv PR)
- percentile rank of the GRE quantitative score (GREq PR)
- percentile rank of the GRE analytic writing score (GREwr PR)
- Biderman's Formula Score that includes uGPA times 100 plus the sum of GREq PR, GREv PR, and GREwr PR (Biderman, 2013)
- the Basic Carnegie Classification from *The Carnegie Classification of Institutions of Higher Education*™ for each institution (The Carnegie Foundation for the Advancement of Teaching, 2010)
- a student's undergraduate institution setting, public versus private
- the Academic Profile of Undergraduate Institutions (APUI)
- whether or not a student took higher level science and math coursework during their undergraduate education

- whether or not the student took advanced athletic training coursework as an undergraduate
- the student's in-state versus out-of-state residency.

Operational Definitions

The following terms are operationally defined for this study.

- Academic Profile of Undergraduate Institution – the best balance between an institution's ACT mean/median and SAT mean/median as a measure of their academic standards
- Adjusted odds ratio – in SPSS is represented by Exp(B), is an indication of a change in the odds one variable has upon the other variables (Field, 2009)
- Biderman's Formula score = (uGPA x 100) + GREv PR + GREq PR + GREwr PR (Biderman, 2013)
- Binary Logistic Regression – a prediction for inclusion into dichotomous categories, natural log rhythm (l_n) times the odds (l_n(odds) (Field, 2009; Peng, Lee, & Ingersoll, 2002; Peng & So, 2002)
- Bivariate - an analysis consisting of two variables, in which neither is identified as an independent (predictor) or dependent (outcome) variables (Mertler & Vannetta, 2005a; Tabachnick & Fidell, 2007)
- Complete data set – for the purposes of this study a candidate had to have the following items as part of their application file:
    - Official copies of transcripts from all colleges and universities attended
    - Official copies of GRE scores

- Confidence interval – a range of potential values for which a population's true values are likely to be contained (Portney & Watkins, 2000)
- Cut-point – also known as a cut-off score, is the score associated with the point on the Receiver Operating Characteristic (ROC) curve which is either closest to the upper left-hand corner or the point furthest away from the diagonal reference line is best determined by Youden's Index
- Exp(B) – Exponent B, is a used in SPSS and is an indication of the adjusted odds ratio
- First-year Graduate Grade Point Average – the GPA for a student at the end of their first year in a GATP
- Higher level science and math coursework – courses established by the GATP which are above the basic level, which may include but are not limited to Biochemistry, Calculus, Histology, Organic Chemistry, Pathophysiology, Physics, and Calculus
- Multicollinearity – occurs when the predictor variables are "very highly correlated ($r \geq 0.80$)" (Mertler & Vannetta, 2005a, p. 342)
- Multivariable – involves the examination of multiple variables (Concato, Feinstein, & Holford, 1993; Feinstein, 1996; Peters, 2008; Reboldi, Angeli, & Verdecchia, 2013; Steyerberg & Harrell, 2003; Tsai, 2013)
- Multivariate – indicates several outcome (dependent) variables (Mertler & Vannetta, 2005a; Peters, 2008; Reboldi et al., 2013)
- Nagelkerke R^2 – analogous to the R^2 in linear regression, a version of the Cox and Snell R^2, provides a measure of the magnitude of the model (Field, 2009)

- Odds Ratio – an estimate of being classified into one category (passing the BOC exam) versus being classified in another category, not passing the BOC exam in case-control studies (Portney & Watkins, 2000); a measure of association which:

$$\frac{p_1}{1-p_1} \Big/ \frac{p_0}{1-p_0}$$

  where p_1 = probability of an event, given the membership in Group 1, p_0 = probability of an event, given the membership in Group 0; an odds ratio of greater than 1.0 implies an increased likelihood; conversely, an odds ratio less than 1 implies a decreased likelihood (Peng et al., 2002; Peng & So, 2002)

- Positive Factor – subject having a score on an predictor variable that is above the established cut-point for the specific predictor variable as established through ROC curve analysis

- Relative Frequency for Success – is similar to Relative Risk,(RR) but since risk is not an appropriate term for this study Relative Frequency for Success (RFS) is being used; is the likelihood that someone who has been classified to be accepted into the GATP will be accepted into the program or is predicted to pass their board exam passes the board exam compared with one who has not been so classified, "indicates the likelihood that someone who has been (classified as meeting the criteria for acceptance will be accepted or to pass the BOC exam will be accepted or will pass the BOC exam), as compared with one who has not (met the criteria to be accepted or to pass the BOC exam)" (adapted from Portney & Watkins, 2000)

- Selection for Admittance into the Graduate Athletic Training Program (GATP) – includes those candidates who have applied to the GATP, have been offered a position in the program regardless of whether the candidate accepted the position and attended classes as part of the GATP. The GATP Selection Committee may select a candidate for admittance, but the candidate may decide to reject position in the program for a variety of reasons
- Success in a GATP – is defined as having a gGPA of greater than or equal to 3.45 at the end of the first year in the GATP
- Undergraduate Grade Point Average – the GPA earned by the subject, is calculated by combining all of the academic institutions a candidate has attended, taken courses and received a grade for academic credit
- Univariable – indicates there is a single predictor variable (Reboldi et al., 2013)
- Univariate – indicates only one outcome variable (Mertler & Vannetta, 2005a; Peters, 2008; Reboldi et al., 2013)
- Youden's Index – is a method to best determine the optimum cut-point on an ROC curve. specifically it is:

$$J = \max_c (Sn(c) + Sp(c) - 1)$$

  - Where:
    - J = Youden's Index
    - c = optimal cut-point for the Sn and Sp - 1
    - $\max_c$ = maximum cut-point on the ROC curve

(Ruopp, Perkins, Whitcomb, & Schisterman, 2008)

Delimitations

The delimitations of this study include admission data from the GATP from 2004 through May 2012 and BOC examination data from 2004 through June 2013. Participants for this study will include those candidates that have applied to the GATP, were offered a position in the GATP, and started the program, students who have fulfilled the academic and clinical requirements for the GATP and are eligible to sit for the BOC exam and take the BOC exam at least one time. For the purpose of creating a prediction model, candidates must pass the BOC examination on their first attempt at taking the exam.

Limitations

The following limitations are acknowledged for this study:

- Effort by candidates on the GRE – candidates have confessed to the authors, that because the GATP does not have a minimum score requirement for the GRE they may not give their best effort on the GRE. Other candidates have confessed they were ill or had other mental and emotional issues that prevented them from giving a better effort on the GRE.
- The undergraduate academic preparation the candidates receive. Each institution, course, and instructor/professor are different in the methods used to evaluate and grade students; therefore, how grades are earned and distributed cannot be controlled, so grade inflation cannot be discounted and prevented.
- The previous clinical experiences the candidate may have prior to their application to the GATP.

- The type of clinical experiences a student in the GATP receives is going to differ for a variety of reasons. These may include, but are not restricted to the following:
    - the location of their clinical rotation
    - the number and kinds of injuries the student may be exposed to
    - the specific preceptor supervising the student and what they are permitted to do or not do under this individual's supervision
    - the number of clinical experience hours which a GATP student earns during their time in the GATP
- Scoring system used to assess the written portion of the GRE is a subjective assessment conducted by a panel of experts.
- Changes that have occurred to the GATP since 2003. These have included but are not restricted to changes in faculty, changes to athletic training competencies and proficiencies, and the teaching responsibilities of the faculty members.
- Whether or not someone is a traditional student. A non-traditional student is defined as someone who delays their enrollment (they do not enter graduate school within [fifteen months] of graduating from their undergraduate school), may be considered financially independent for financial aid purposes, has dependents, or is a single parent (modified from the definition provided by Horn & Carroll, 1996).

Assumptions

The following assumptions are made for this study:

- The percentile rank related to the old and new GRE scoring system as provided by Educational Testing Services are accurate (Educational Testing Services, 2013a, 2013b)
- The first-time certification data provided by the BOC are accurate (Board of Certification (BOC) Certification Examination for Athletic Trainers, 2008, 2009; Johnson, 2010, 2011, 2012, 2013; National Athletic Trainers Association Board of Certification, 2005, 2006; National Athletic Trainers' Association Board of Certification, 2003, 2004; National Athletic Trainers' Association Board of Certification Inc., 2002, 2007).
- That the statement made on the University of Tennessee at Chattanooga Psychology Department's web site is accurate when the Department reports Biderman's Formula Score "has been found to be significantly related to performance in the program" (Biderman, 2013).
- The information provided by a university's web site related to their common data set (Common Data Set Initiative, 2012) is accurate.
- The information provided by *The Carnegie Classification of Institutions of Higher Education*™ (The Carnegie Foundation for the Advancement of Teaching, 2010) is accurate.

Summary of Chapter

Chapter I provided a brief synopsis of this study. This study had two interrelated purposes, both of which pertained to the process of admitting students to a graduate professional program. The first component of this study involves the development of a prediction model to identify factors associated with eligibility and first-attempt success on the Board of Certification

(BOC) examination for students who have completed a professional (entry-level) graduate athletic training program (GATP). The second component will utilize the results of the first analysis to identify program applicant characteristics that are most likely to predict both academic success within the graduate professional program and subsequent success on the BOC exam. This chapter outlined the statement of the problem, hypotheses, dependent and predictor variables, operational definitions for the study, delimitation, limitations, and assumptions that are anticipated at this point in the dissertation process.

Chapter II

LITERATURE REVIEW

Introduction

Many health education program administrators have examined potential predictors for assisting in their decisions to admit or reject students. A literature search on programs from clinical psychology (Daehnert & Carter, 1987), nursing (Hansen & Pozehl, 1995; Katz, Chow, Motzer, & Woods, 2009; Munro, 1985; Newton & Moore, 2007; Salvatori, 2001; Wilson, 1999), occupational therapy (Kirchner & Holm, 1997; Salvatori, 2001), physician assistant (Hocking & Piepenbrock, 2010), physical therapy (Balogun et al., 1986; Kirchner et al., 1994; Levine et al., 1986; McGinnis, 1984; Morris & Farmer, 1999; Payton, 1997; Zipp et al., 2010), respiratory care (Salvatori, 2001), and medical school (Ferguson et al., 2002; Meleca, 1995; Salvatori, 2001) find that all have attempted to refine their selection processes. Several different variables have been used in the hope of trying to either isolate or find which group of variables may provide the best prediction model to determine the candidates that should be selected for their programs. Predictor variables such as the Graduate Record Examination (GRE) (Daehnert & Carter, 1987; Hocking & Piepenbrock, 2010; Katz et al., 2009; Kirchner & Holm, 1997; Munro, 1985; Newton & Moore, 2007), undergraduate grade point average (uGPA) (Daehnert & Carter, 1987; Hansen & Pozehl, 1995; Hayes et al., 1997; Keskula et al., 1995; Kirchner & Holm, 1997; McGinnis, 1984; Meleca, 1995; Munro, 1985; Newton & Moore, 2007; Salvatori, 2001; Silver & Hodgson, 1997), Medical College Admission Tests (MCAT) (Kreiter & Kreiter, 2007; Meleca, 1995; Salvatori, 2001; Silver & Hodgson, 1997), past clinical experience (Ferguson et al., 2002; Hansen & Pozehl, 1995; Hayes et al., 1997), age (Hansen & Pozehl, 1995), race, gender (Ferguson et al., 2002), and ethnicity (Ferguson et al., 2002) have all been used. Several other more subjective variables have also been used to measure successful candidates and have included written essays, interviews, subjective inventories, references, and personal characteristics (Balogun et al., 1986; Bretz, 1989; Burton & Wang, 2005; Day, 1986; DeAngelis, 2003; Feldman, 2007; Ferguson et al., 2002; Hayes et al., 1997; Keskula et al., 1995; Kirchner & Holm, 1997; Kirchner et al., 1994; McGinnis, 1984; Meleca, 1995; Mitchell, 1990; Munro, 1985; Payton, 1997; Platt et al., 2001; Sime et al., 1983; Utzman et al., 2007a; Willingham, 1972; Zipp et al., 2010). The outcome variables that have been used include admission into the program, graduate grade point average (gGPA), academic difficulty, academic performance, clinical rotation success, and graduation from the program (Balogun et al., 1986; Bretz, 1989; Burton & Wang, 2005; Day, 1986; DeAngelis, 2003; Feldman, 2007; Ferguson et al., 2002; Hayes et al., 1997; Keskula et al., 1995; Kirchner & Holm, 1997; Kirchner et al., 1994; McGinnis, 1984; Meleca, 1995; Mitchell, 1990; Munro, 1985; Payton, 1997; Platt et al., 2001; Sime et al., 1983; Utzman et al., 2007a; Willingham, 1972; Zipp et al., 2010).

There are currently no studies that have examined admittance decisions for professional (entry-level) graduate athletic training programs (GATP). Keskula, Sammarone & Perrin (1995) studied prediction variables for post-professional NATA-approved Graduate Athletic Training Programs. They used stepwise multiple regression analysis to determine uGPA was the only significant predictor of gGPA (Keskula et al., 1995).

Medical professions have a board certification or licensure examination process to pass before being eligible to practice their profession. Graduates become eligible to sit for these accrediting exams upon completion of their education. The primary purpose of these exams is to determine the entry-level competence of the candidate and to protect the health and welfare of the general public (Federation of State Boards of Physical Therapy, 2012; National Athletic Trainers' Association Board of Certification Inc., 2006; National Board for Certification in Occupational Therapy, 2009; United States Medical Licensing Examination, 2012). Several professions or medical specialties such as: medicine (Ferguson et al., 2002), nurse anesthetists (Zaglaniczny, 1992), obstetrics and gynecology (Armstrong et al., 1998), physical therapy (Kosmahl, 2005; Utzman et al., 2007a), and surgery (de Virgilio et al., 2010), have tried to create their own prediction models for passing their certification/licensure exams with varied success. Predicting achievement on the BOC exam has been limited (Erickson & Martin, 2000; Harrelson et al., 1997; Hickman, 2010; Middlemas et al., 2001; Pickard, 2003; Starkey & Henderson, 1995; Turocy et al., 2000; Williams & Hadfield, 2003). Therefore, the second purpose of this review was to examine the ability of health related professions to predict success on their certification or licensure exams.

This review will begin with a brief a history of athletic training education and the BOC examination. A discussion of prediction modeling will also be included in this review.

History of Athletic Training Education

Athletic training's birth likely occurred in ancient Greece with the creation of the Olympics (Ebel, 1999). In the United States, Harvard hired James Robinson as the first athletic trainer in 1881. In 1932 a group of athletic trainers were present at the Summer Olympics Games in Los Angeles, CA (Ebel, 1999).

In the 1950s, a group of about 200 athletic trainers met in Kansas City and formed the National Athletic Trainers' Association (Ebel, 1999; National Athletic Trainers' Association, 2011a). By 1959 recommendations for educational requirements in the colleges and universities was proposed; however, ten years later only four colleges/universities had established athletic training educational programs. In 1973 there were 14 colleges/universities with approved undergraduate athletic training curriculum programs and by 1978, 46 colleges/universities had approved undergraduate athletic training programs (Delforge & Behnke, 1999; Ebel, 1999; Lindquist, Arrington, & Scheopner, 2007).

From 1969 until 2004 there were two routes to qualify to sit for the BOC exam. A student could graduate from an approved athletic training professional (entry-level) education program (undergraduate or graduate) or through an apprenticeship/internship program with a bachelor's degree (Delforge & Behnke, 1999; Ebel, 1999; Lindquist et al., 2007). The internship route to certification was terminated in 2004 (Lindquist et al., 2007). Presently there are over 350 professional (entry-level) undergraduate athletic training programs (National Athletic Trainers' Association, 2011b) and 27 professional (entry-level) graduate athletic training programs (Commission on Accreditation of Athletic Training Education, 2013d). How students have been accepted into a school's athletic training educational program has varied from school-to-school. The only admission requirements mandated by athletic training's accrediting body, the Commission on Accreditation of Athletic Training Education, are athletic training education programs must be in compliance with the Americans with Disabilities Act of 1990 (Harkin, 1990; National Athletic Trainers' Association, 2000).

History of BOC Exam

In the spring of 1969, J. Lindsy McLean wrote an article for the *Journal of the National Athletic Trainers Association* (a predecessor to the present *Journal of Athletic Training*) asking whether or not the NATA needed a certification exam (reprinted in 1999). By December 1969, the NATA had implemented a process for becoming a certified athletic trainer. In August 1970, the first certification examination was administered (Grace, 1999; Lindquist et al., 2007). The initial exam had two portions, a written section which contained 150 multiple choice questions and three oral-practical exam questions. By June 1985 a written simulation portion was added to the certification exam. The written simulation portion presented students with scenarios and asked the student what steps they would take as they worked their way through the situation. In order for a student to become certified he or she had to have graduated from an accredited athletic training education program and have passed all three portions of the certification exam (Lindquist et al., 2007).

In 1995, the oral-practical section of the exam became an assessment of psychomotor skills only as the oral portion of the exam was dropped. The psychomotor assessment portion of the exam was eventually discontinued after the April 2007 exam date. By June of that year the entire exam was computerized (Lindquist et al., 2007).

The BOC created the certification examination to determine the competency of athletic training students. Questions for the BOC examination are developed by a committee of certified athletic trainers. The questions are created based on the *BOC Role Delineation/Practice Analysis*, which is broken into eight main content areas (Board of Certification, 2011a). The eight content areas or domains of athletic training are:

1. Evidence-based Practice
2. Prevention and Health Promotion
3. Clinical Examination and Diagnosis
4. Acute Care of Injury and Illness
5. Therapeutic Interventions
6. Psychosocial Strategies and Referral
7. Healthcare Administration
8. Professional Development and Responsibility

(Board of Certification, 2011a, 2011b; National Athletic Trainers' Association, 2011).

Once an exam question is created it is then submitted to group of independent evaluators for the questions to be validated. Questions are cross referenced from the literature, edited for grammar, content, technical adequacy and clarity. If a question is deemed to be appropriate, then it may be placed on the exam as an experimental/unscored item. Based on the evaluation process these experimental questions are then appraised for future use on the BOC exam or the need for further revision and assessment (Board of Certification, 2011a).

In 2011, the BOC exam consists of 175 questions and candidates have four hours to complete it. Only 150 questions are used for the scoring portion of the exam, while the remaining questions are the “test” or experimental questions for potential inclusion in future exams. Although all questions are scored, only those questions which are not test/experimental questions are applied to the candidate's exam performance for passing or not passing the certification exam. The candidate does not know which questions are to be scored as part of the actual exam or which questions are experimental (Board of Certification, 2011a).

The BOC exam questions are of three different types:

1. Stand-alone multiple-choice questions
2. Stand-alone alternative items (drag-and-drop, text based simulation, multi-select, hot spot, etc.)
3. Focused testlets
    a. A 5-item focused testlet consists of a scenario followed by 5 key/critical questions related to that scenario
    b. Each focused testlet may include multiple-choice questions or any of the previously described alternative item types (Board of Certification, 2011a"Development: Format," para. 1)

The passing point for the BOC exam is established through the use of the Angoff method (Board of Certification, 2011a), which uses a "panel of judges" to "examine each multiple-choice item" and "estimates the probability that the 'minimally competent' candidate would answer the item correctly" (George, Haque, & Oyebode, 2006, p. 47). The mean of the probabilities is then calculated and this determines the passing point for the BOC exam. Reliabilities are computed for each of the domains of athletic training. For each new exam, the passing point and reliabilities are calculated back to the initial version of the exam to assure fairness to the candidates so the specific test an individual is taking is not significantly easier or harder than taking a different variation of the exam (Board of Certification, 2011a; George et al., 2006)

The first time success rate on the BOC exam has varied through the years. Williams and Hadfield (2003), reported that the first time pass rate for all three section of the exam from 1997-2002 was only 35%. From the BOC testing year of 1995-1996 through the 2011-2012 exam year, the success rate for first-time candidates passing the BOC exam has varied from 30-82% with an overall average during this time of 47.9% and a median of 48.4% (Board of Certification (BOC) Certification Examination for Athletic Trainers, 2008, 2009; CASTLE Worldwide, 2001; Henderson, 1998; Johnson, 2010, 2011, 2012, 2013; National Athletic Trainers Association Board of Certification, 2005, 2006; National Athletic Trainers' Association Board of Certification, 2003, 2004; National Athletic Trainers' Association Board of Certification Inc., 1997, 1999, 2000, 2002, 2007). The overall average from the 1995-1996 exam years through the most recent report, 2012-2013 exam year, is 49.7%. Figure 2.1 provide the data for the year-by-year first-time pass rates.

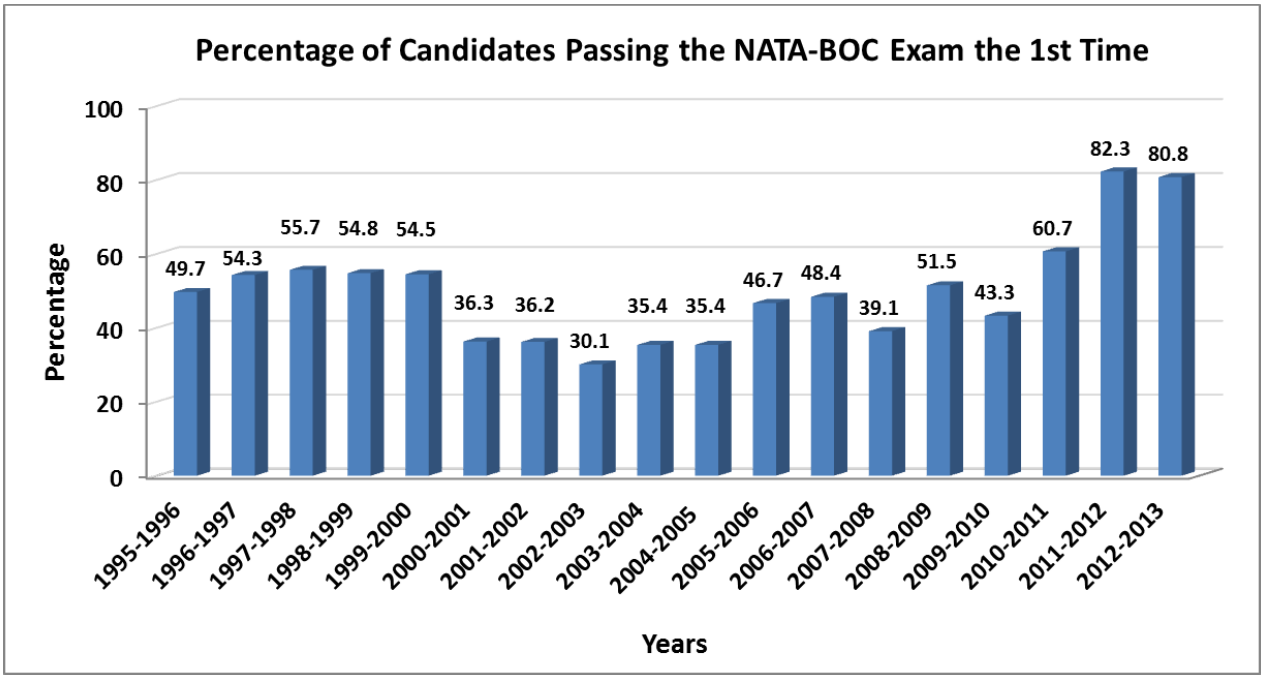

Figure 2.1 First time pass rates by examination year

Admission to Health Related Programs

Several graduate level, allied health programs have examined potential predictors to assist in the admittance decisions to their programs and to assist in determining which candidates might have a better opportunity at success. Programs in clinical psychology (Stricker & Huber, 1967), nursing, (Hansen & Pozehl, 1995; Munro, 1985; Newton & Moore, 2007; Rhodes, Bullough, & Fulton, 1994) occupational therapy (Kirchner & Holm, 1997), physician assistant (Hayes et al., 1997), physical therapy (Day, 1986; Kirchner et al., 1994; Levine et al., 1986; Rhodes et al., 1994; Zipp et al., 2010) and medical schools (Meleca, 1995; Mitchell, 1990) have had varying degrees of success in their ability to select potential candidates for their programs to determine whom might have a better opportunity at success. The predictor variables used included: Graduate Record Examination, (GRE), (Day, 1986; Hansen & Pozehl, 1995; Hocking & Piepenbrock, 2010; Kirchner & Holm, 1997; Kirchner et al., 1994; Munro, 1985; Newton & Moore, 2007; Rhodes et al., 1994; Thacker & Williams, 1974), undergraduate grade point average, (uGPA), (Hansen & Pozehl, 1995; Julian, 2005; Kirchner & Holm, 1997; Kirchner et al., 1994; Levine et al., 1986; Meleca, 1995; Mitchell, 1990; Munro, 1985; Newton & Moore, 2007; Salvatori, 2001; Stricker & Huber, 1967; Templeton, Burcham, & Franck, 1994; Thacker & Williams, 1974; Utzman et al., 2007a; Zipp et al., 2010), Medical College Admission Tests (MCAT), (Julian, 2005; Meleca, 1995; Mitchell, 1990; Salvatori, 2001), age (Hayes et al., 1997; Utzman et al., 2007a), race (Utzman et al., 2007a), gender (Hayes et al., 1997), and ethnicity (Utzman et al., 2007a). Some of these same studies have indicated more subjective measures to predict success have been utilized. These have consisted of: written essays (Kirchner & Holm, 1997; Munro, 1985), interviews (Hayes et al., 1997; Levine et al., 1986), the Problem Solving Inventory (DeAngelis, 2003), references (Kirchner & Holm, 1997), and personal characteristics (Levine et al., 1986). A variety of outcome variables have been examined to predict success to include: admission into the program, graduate grade point average (gGPA), academic difficulty (Utzman et al., 2007a), academic performance (DeAngelis, 2003; Hayes et al., 1997; Julian, 2005; Kirchner & Holm, 1997; Kirchner et al., 1994; Stricker & Huber, 1967; Thacker & Williams, 1974; Zipp et al., 2010), clinical success (Kirchner & Holm, 1997; Munro, 1985), and graduation verses non-graduation from the program (Williams, Harlow, & Stable, 1970). None of the studies could specifically define success for their specific programs or professions. No studies to date have examined admission predictors of graduate athletic training education programs.

Prediction for Success on Certification/Licensure Exams

Attempts to predict success on national examinations for certification/licensure examinations have been made by several professions or medical specialties. Anesthesiology (McClintock & Gravlee, 2010), athletic training (Draper, 1989; Erickson & Martin, 2000; Harrelson et al., 1997; Hickman, 2010; Middlemas et al., 2001; Pickard, 2003; Starkey & Henderson, 1995; Turocy et al., 2000; Williams & Hadfield, 2003) medicine (Armstrong et al., 1998; Johnson, 2010), and physical therapy (Utzman et al., 2007a; Vendrely, 2007) have attempted to predict success on their board or licensing examinations to varying degrees. Predictor variables used to assist in these prediction models include:

- Anesthesiology
    - Country of medical school
    - Gender
    - In-Training Examination (ITE) scores
    - Residency program accreditation cycle length (McClintock & Gravlee, 2010)
- Medicine
    - American Board of Surgery (ABS) In-Training Examination (ABSITE) score
    - Fellowship training
    - Mandatory research
    - Residency program type
    - Surgical volume (Johnson, 2010)

- U. S. Medical Licensure Examination step 1 and grade point average (Armstrong et al., 1998)

- Physical therapy
  - California Critical Thinking Skills Test (CCTST)
  - Demographic characteristics (race, age and ethnicity)
  - Final GPA
  - GREq (Vendrely, 2007)
  - GREv (Utzman et al., 2007a)
  - Ratings on the Clinical Performance Instrument (CPI)
  - uGPA (Vendrely, 2007)

In athletic training nine different studies have been published which have attempted to predict success on the Board of Certification (BOC) exam, (Draper, 1989; Erickson & Martin, 2000; Harrelson et al., 1997; Hickman, 2010; Middlemas et al., 2001; Pickard, 2003; Starkey & Henderson, 1995; Turocy et al., 2000; Williams & Hadfield, 2003). There was a variety of predictor variables used in these athletic training studies. Student learning style was used by Draper (1989), uGPA including overall GPA, athletic training GPA and academic minor GPAs (Harrelson et al., 1997; Middlemas et al., 2001), type of athletic training preparation program the student came from, either an accredited curriculum program or the internship route (Middlemas et al., 2001; Starkey & Henderson, 1995), “ACT composite score, and the number of semesters of university enrollment” (Harrelson et al., 1997, p. 327), the number of clinical experience hours, previous athletic training experience and demographic data (Hickman, 2010; Middlemas et al., 2001; Turocy et al., 2000), the students’ football experience (Hickman, 2010), and the academic year the athletic training student began their clinical rotations (Hickman, 2010) have all been examined.

Turocy (2002) states two studies (Harrelson et al., 1997; Middlemas et al., 2001) found that uGPA was the strongest predictor. However, we find two flaws in her assessment (Turocy, 2002).

In the study by Middlemas et al. (2001) they found consistent predictors for success on the exam as a whole (passing all three portions of the BOC exam) in GPA, clinical hours completed and route to the certification exam (accredited curriculum versus internship). They did not find any predictors for any single portion of the exam, but Middlemas et al. (2001) we believe the R^2 value is too small to draw any substantial conclusions ($R^2 = 0.057$).

Harrelson et al. (1997) found academic performance to be a strong predictor for first-time success on the BOC exam on all three sections of the exam. However, they only had 52 subjects in their study and the authors acknowledge there were problems in their study with the predictive power of their independent variables. They used overall GPA, “athletic training GPA, academic minor GPA, ACT composite score, and number of semesters enrolled at (their) university” (p. 324). Although they were able to account for a more meaningful degree of the variance accounted for the entire examination ($R^2 = 0.26$), when considering each of the individual sections of the exam, their R^2 values are not strong: written portion of the exam ($R^2 = 0.12$); written simulation portion ($R^2 = 0.11$); oral/practical section ($R^2 = 0.10$). The authors also did not mention what their effect size was or report the confidence intervals related to their data (Harrelson et al., 1997).

Erikson and Martin (2000) had success in their Delphi study to predict athletic training student success on the BOC exam. They used a panel 35 experts who identified 66 items they perceived as contributing factors to first-time success on the BOC exam. Some of these factors included "ability to interpret the question" (p. 135), "knowledge of theories and techniques in rehabilitation and modalities" (p. 136), "clinical settings that allow students to take an active role" (p. 136),"instructors committed to providing a positive learning environment" (pp. 136-137), and "clinical assessment skills" (p. 137). Unfortunately, the study did not test these attributes on actual candidates taking the BOC exam for the first time; nor was there a follow-up study done to examine the reliability or validity of these predictors.

A common factor for all nine of the athletic training studies in their attempt to predict success on the BOC exam is that the data were gathered using the performance of undergraduate students (Draper, 1989; Erickson & Martin, 2000; Harrelson et al., 1997; Hickman, 2010; Middlemas et al., 2001; Pickard, 2003; Starkey & Henderson, 1995; Turocy et al., 2000; Williams & Hadfield, 2003). Currently there have been no studies conducted that have examined potential prediction variables for the success of students from professional (entry-level) GATPs and the success they have had on BOC exam.

The national first time pass rate on the BOC since 2007 through 2011 was 48.6% (Johnson, 2012), (Figure 2.1) while the first time pass rate on the BOC over the same time period for students from the GATP is 83.2% (Bruce, 2011). The purpose of this study is to create a prediction model to estimate success on the Board of Certification exam by students coming from a professional (entry-level) graduate athletic training program.

Statistical Analysis

Frequentist Statistics versus Bayesian Philosophy

There are two main statistical schools of thought: frequentist and Bayesian. Both methods explore probability, but the theories and the methods are very different (Vallverdú, 2008). The Bayesian approach to probability is to "measure the degree of belief in an event, given the information available." The focus in on the individual's "state of knowledge" rather than a "sequence of events" (Vallverdú, 2008, Bayesian approach section, para. 1). The frequentist approach to probability interprets it as "a long-run frequency of a 'repeatable' event." With a frequentist's approach "probability would be a measureable frequency of events determined from repeated experiments" (Vallverdú, 2008, Frequentist approach section, para. 1).

The Bayesian approach to statistics originated in England by a minister named Thomas Bayes when it was first described in an article in 1763. The paper, submitted posthumously, described what became known as the Bayesian theorem in which the estimated probability of an event occurring or being true, the estimated probability of an event not occurring or being false, and the third is to estimate the prior probability (or simply known as a prior). A prior is defined as the probability you would assign to an event of occurring before you received additional information (Silver, 2012). "The most practical definition of a Bayesian prior might simply be the odds at which you are willing to place a bet" (Silver, 2012, pp. 255-256). Bayesian's priors can remain strong and resilient even when there is new information (Silver, 2012). The efficiency and effectiveness of using prior or historical information will enhance many statistical models (Rothman, Greenland, & Lash, 2008; Silver, 2012). Algebraically the Bayesian theorem is demonstrated in Table 2.1:

Table 2.1 Equation for Bayes' Theorem

$$\text{Bayes' Theorem} = \frac{XY}{XY + Z\,(1 - X)}$$

Where: X = prior probability
Y = probability of event occurring or being true
Z = probability of event not occurring or being false (Silver, 2012)

Another way of looking at Bayes' Theorem is to understand it as "a relationship of probabilities and 'conditional' probabilities" (Hubbard, 2010, p. 178). A conditional probability is characterized as "the chance of something given a particular condition" will or will not occur (Hubbard, 2010, pp. 178-179). Table 2.2 demonstrates this form of the Bayes' Theorem:

Table 2.2 Equation for Bayes' Theorem for Probabilities

$$P(A \mid B) = \frac{P\,(A) \times P\,(B \mid A)}{P\,(B)}$$

Where: P (A | B) = conditional probability of A given B
P (A) = probability of A
P (B) = probability of B
P (B | A) = conditional probability of B given A

The major rival to Bayesian philosophy came from another Englishman, Ronald Aylmer (R. A.) Fisher, who was born about 120 years after Bayes died. Fisher is the individual who developed many of the statistical methods still used today. His creation of statistical significance and the associated methodology focused on helping the data to be freer of bias or contamination. The focus of Fisher's techniques relies on selecting a representative sample from a population and applying the results from the sample to the population. This form of statistics became known as frequentism or frequentist (Silver, 2012).

The frequentist approach to statistics has been the dominate form of statistics in research since the 1920s. The ideas conveyed by R. A. Fisher, John Venn, Jerzy Neyman, and Egon Pearson caused researchers to shift their paradigm. The concepts they discussed and advocated for were that of relative frequency. The researcher would perform the experiment many times and the count the number of subjects who achieved or had a positive outcome or result (Vallverdú, 2008; Zabell, 1989).

Advocates of frequentist statistics criticize the Bayesian approach as being overly subjective and arbitrary. Bayesians defend this by stating there is an element of subjectivity and arbitrary elements in all statistical inferences (Rothman et al., 2008). Frequentist seek to avoid the reasons behind why predictions most often are wrong, that being human error. Bayesian philosophy helps to apply problems into a the real world, while frequentist statistics are more confined to the laboratory and less suitable for the real world (Rothman et al., 2008; Silver, 2012).

From the 1760s into the 20th Century the Bayesian approach was the dominate statistical technique (Fienberg, 2006) (Tables 2.3 and 2.4 demonstrate the differences between the two forms of statistical analyses.) The label Bayesian did not come into the lexicon until 1970s. Thomas Bayes created many of the methods and theories used in probability testing with its roots associated with “inverse probability”. The term inverse was used because “it involves inferring backwards from the data to the parameter or from effects to causes” and led to what is known today as inferential statistics (Fienberg, 2006, p. 5).

From about 1950 into the 1990s almost no one utilized Bayesian philosophy, except for a few researchers on the fringe of science. There were a couple of reasons for this occurrence. First, everybody was engaged in the cookbook mentality of using a certain frequentist procedure if a specific type of study was being performed. Researchers were led to believe that frequentist statistics was the way things had always been done and it was the most popular form of statistical analysis (Casella, 2008). We conducted a brief search of academic search engines using just the terms Bayesian and frequentist. Tables 2.3 and 2.4 demonstrate how inaccurate this assumption was and continues to be today. Table 2.3 shows all of the references without any restrictions on the dates. Table 2.4 shows the references from the 21st Century.

Table 2.3 Comparison of the Number of References on Various Academic Search Engines

Search Engine	**Bayesian**	**Frequentist**
Google Scholar	1,070,000	29,300
WorldWideScience.org	368,736	39,746
Science.gov	209,119	36,951
Microsoft Academic Search	119,288	2,937
PubMed	18,639	506
The Cochran Library	7,646	680,109
Library of Congress	128	1
Digital Library of the Commons	96	6
Total	**1,793,652**	**789,556**

Table 2.4 Comparison of the Number of References in the 21st Century on Various Academic Search Engines

Search Engine	Bayesian	Frequentist
Google Scholar	580,000	19,700
WorldWideScience.org	331,298	31,224
Microsoft Academic Search	88,932	2,136
PubMed	16,635	454
Total	**1,016,865**	**53,514**

Secondly, computers were very cumbersome, slow, and unavailable to the masses during the first 90 years of the 20th Century. It was not until the 1990s when personal or desk top computers became much more affordable and easier to use. The third reason relates closely to the second – Bayesian philosophy takes a lot of computation and to do these computations in long hand takes a great deal of time and increases the risk of error (Casella, 2008).

In a frequentist's world, the data are generated by repeating the experiment on a random sample (providing the frequency of an event), the basic limitations remain the same during the application of the repeatable experiment; therefore the parameters are constant. In the Bayesian's world the data are gathered from an observed cohort, the parameters are unspecified and are described in terms of the likelihood of an event occurring or not occurring; therefore the data are fixed (Casella, 2008). Bayesian philosophy is about observing the "association between the exposure and the outcome" (Denegar & Wilkerson, 2013, slide 27). For the purposes of this study the exposures are the traits (predictor variables) students possess. The outcomes are either being accepted or not being accepted into the GATP or passing or not passing the BOC exam the first time a student takes the exam.

Univariate/Multivariate vs. Univariable/Multivariable

Throughout the literature, especially medical literature, researchers tend to disagree on the appropriate use of the terms univariate / univariable, multivariate / multivariable. Although the use of these terms is often used interchangeably they have very different meanings and connotations (Concato et al., 1993; Reboldi et al., 2013). The suffix –variate refers to the outcome or dependent variable (Concato et al., 1993; Feinstein, 1996; Mertler & Vannetta, 2005a; Peters, 2008). The term variate is defined as “a random variable with a numerical value that is defined on a given sample space” (The Free Dictionary by Farlex, 2000c, variate) The suffix –variable refers to the predictor or independent variable (Concato et al., 1993; Peters, 2008; Steyerberg & Harrell, 2003). The term variable refers to “having no fixed quantitative value” or the capability “of assuming any . . . set of values” (The Free Dictionary by Farlex, 2000b, variable).

Univariable analysis in where there is a single predictor variable. This form of analysis is often used in the determination of the inclusion or exclusion of variables based on some sort of criteria (Reboldi et al., 2013). Multivariable analysis involves multiple predictor variables (Concato et al., 1993; Feinstein, 1996; Peters, 2008; Reboldi et al., 2013; Steyerberg & Harrell, 2003; Tsai, 2013). There are three general models in which multiple variables can relate to one another. They are as follows:

1. Multiple predictor (independent) variables relating to a single outcome (dependent) variable, known as a “many-to-one relationship.”
2. Multiple predictor (independent) variables relating to multiple outcome (dependent) variables, known as “many-to-many relationship.”

3. Multiple variables which are neither predictor or independent nor outcome or dependent. This is known as a "many-internal relationship, being interrelated to one another, but not to the external variable." (Feinstein, 1996, p. 2)

Any of these three forms of relationships are referred to as multivariable since there is more than one predictor variable. This was further reinforced by J. Concato (personal communication, December 12, 2013) when he stated if there is more than one predictor variable, the term multivariable is warranted.

The main statistical methods utilized in multivariable analysis differ "in the expression and format of the outcome expressed as the dependent variable (Concato et al., 1993, p. 201). These methods include:

1. Multiple linear regression has a continuous outcome variable
2. Multiple logistic regression has a binary outcome variable and "occurs at a fixed point in time."
3. Discriminant function analysis has an outcome variable which the subject belongs to a category or a group where there are more than two possible outcomes.
4. Cox regression has an "outcome variable which is duration of time to occurrence of a binary 'failure' event during a follow-up period of observation." Simply put what is the subject's outcome status at the time when the study is terminated (Concato et al., 1993, pp. 201-202).

The use of terms univariate, bivariate, and multivariate often are used without regard to what they actually signify (Feinstein, 1996). Univariate refers to a single outcome variable although there may be many predictor variables (Peters, 2008; Tabachnick & Fidell, 2007).

When an analysis consist of two variables, but neither is identified as either an independent or dependent variables it is known as bivariate (Mertler & Vannetta, 2005a; Tabachnick & Fidell, 2007). Multivariate analysis indicates several outcome (dependent) variables, while when there is only one outcome variable the proper term is univariate (Mertler & Vannetta, 2005a; Peters, 2008; Reboldi et al., 2013). Most common medical studies are not multivariate since there is usually one outcome variable (Concato et al., 1993). Conversely, Tabachnick & Fidell, (2007), state that multivariate analysis includes simultaneous analysis of multiple outcome and multiple predictor variables.

For our study, since the development of the prediction models is similar the terms univariable, multivariable and univariate will be used. We are examine each predictor variable individually (univariable), then combine the variables for further investigation (i.e., multiple predictor variables; therefore, multivariable), and a single outcome variable (univariate).

Evidence-based Research

Evidence-based research came out of the practice of evidence-based medicine (EBM). By definition EBM is "the integration of the best research evidence" with clinical expertise and the patient's unique values and circumstances (Straus, Richardson, Glasziou, & Haynes, 2005, p. 1). Evidence-based medicine has become multidisciplinary for a variety of allied health care professions and the evolving research has enabled practitioners to share and communicate related information. It allows the clinician to seek and access answers to questions and incorporate the information into effective therapies and interventions. Evidence-based medicine also allows the clinician to focus their reading on the specific issues that arise in their clinical practice rather than randomly seeking answers through the mass of literature and myriad of journals and that are available today (Sackett, 1997; Steves & Hootman, 2004).

Figure 2.2 represents the early model of what was involved in EBM. It was the combination of clinical expertise along with the best research evidence available and the preferences and values of the patient (Haynes, Devereaux, & Guyatt, 2002).

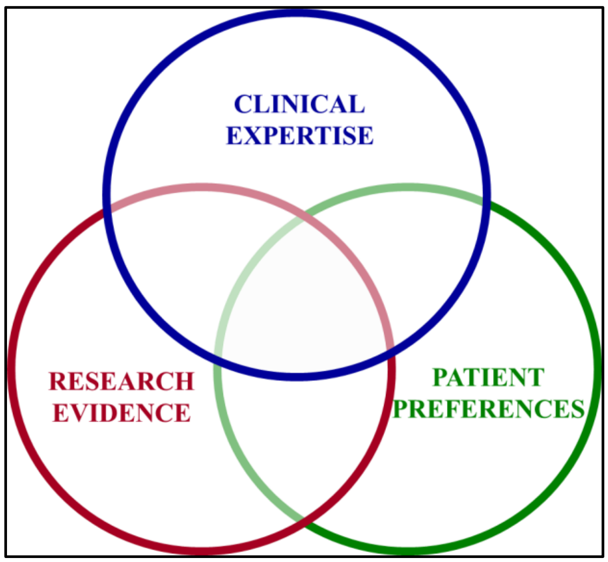

Figure 2.2 Early Model for Evidence-based Medicine

Figure 2.3 shows an updated model of EBM where clinical presentation of the patient along with the best available research evidence available and the preferences and values of the patient are all considered as part of the expertise of the clinician to provide the best possible care available (Haynes et al., 2002).

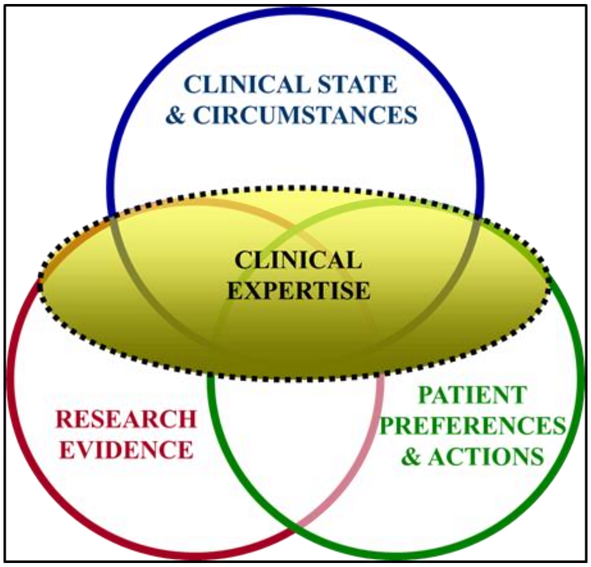

Figure 2.3 Updated Model for Evidence-based Medicine

According to the Oxford Centre for Evidence-based Medicine, the ranking of the levels of evidence is listed on five different levels. The gold standard of research is the randomized controlled trial (RCT), Level I. Level II are cohort studies while case-control studies are classified as Level III studies. A case-series is a Level IV study, and expert opinion is seen as the weakest and is classified at Level V (Phillips et al., 2009). Frequentist advocates tend to stay at Level I and do not give much credence to cohort studies. The problem with the RCT is conducted in a relatively sterile environment where all of the extraneous variables are controlled as best as possible to determine if only one variable is responsible for the change in condition (Portney & Watkins, 2000). Although the ultimate goal would be to strive to conduct the highest level of research possible (Level I), it is not always practical to control all of the variables, nor randomize all of the subjects.

A cohort study (Level II) can provide valuable information when desiring to follow a like group of individuals over a specific period of time. In clinical prediction modeling research, the clinician acknowledges it is impossible to control for all of the extraneous variables of our patients. These studies have great utility due to the larger populations compared to the relatively small number of subjects utilized in RCTs. Cohort studies are also advantageous in their ability to classify subjects in any one of many different categories depending upon the study. Many epidemiological studies examine the presences or absence of some sort of condition, disease, or injury and related to some exposure. This technique permits the researcher to easily classify the patients or subjects into one of four general categories, which provides for a 2 X 2 contingency table to be established. From this 2 X 2 table any number of different statistical procedures can be calculated to provide a variety of data (Denegar & Cordova, 2012; MacDermid & Law, 2007; Portney & Watkins, 2000).

The United Kingdom's National Health Services (NHS) expands the Oxford Centre's rating system whereby their top rating is: 1++; 1+; 1; 2++; 2+; 2; 3; 4. Their "1" raking include meta-analyses and systematic reviews of RCTs. A 1++ has very little risk of bias; a 1+ has low risk of bias, while a 1 has a high risk of bias. The NHS rating of 2 studies other than RCTs, so a 2++ is a systematic review of a cohort or case-control study. A 2+ rated study is a well-designed and conducted cohort or case-control studies that possess few confounding bias variables. If a cohort or case-control study has a high risk of having confounding variables then it is rated as a 2 by the NHS system. Studies rated as a 3 are case series or case report studies, while a rating of 4 is expert opinion (MacDermid & Law, 2007)

From EBM the evolution of clinical prediction rules were developed. There are three main steps when creating a prediction rule: 1. derivation of the prediction rule; 2. validation of the rule; 3. impact analysis (Bruce & Wilkerson, 2010a; Childs & Cleland, 2006). Creating the prediction rule is to identify all of the potential predictors (Bruce & Wilkerson, 2010a; Childs & Cleland, 2006) or “priors” according to Silver (2012). Validating the rule is to apply the rule to a different population of similar patients/subjects for which the rule was originally created. The final step is to conduct an impact analysis. This step involves evaluating whether or not the rule affected the clinician’s behavior, improved outcomes or reduced costs (Bruce & Wilkerson, 2010a; Childs & Cleland, 2006). Bayesian statistics not only helps to create the prediction rule, but assists in the determination of the accuracy of a prediction in the real world (Silver, 2012).

Prediction Modeling

Prediction models have been used in medicine to guide clinical practice for some time (Beneciuk, Bishop, & George, 2009; Bruce & Wilkerson, 2010b; Childs & Cleland, 2006; Emparanza & Aginaga, 2001; Flynn et al., 2002; Heyworth, 2003; Kuijpers et al., 2006; Laupacis, Sekar, & Stiell, 1997; Stiell, 1996; Stiell et al., 1992 ; Wasson, Sox, Neff, & Goldman, 1985). These prediction models have been called clinical prediction rules or clinical decision rules (Bruce & Wilkerson, 2010a; Childs & Cleland, 2006; Childs et al., 2004; Childs, Fritz, Piva, & Erhard, 2003; Cleland, Childs, Fritz, Whitman, & Eberbart, 2006; Cleland, Childs, Fritz, Whitman, & Eberhart, 2007; Haswell, Gilmour, & Moore, 2008; Laslett, 2006; Podichetty & Morisue, 2009; Yealy & Auble, 2003). An argument could be made that “clinical prediction guide” is a better term since rules are usually hard and fast (Denegar, 2012). Breaking rules usually results in consequences. While guides assist someone in negotiating territory and there are few if any consequences when they are violated (Bruce, 2012; Denegar, 2012). Regardless of what they are called their primary purpose is to aid the healthcare practitioner in making clinical decisions about a condition or how to treat the specific condition based on the research evidence.

Prediction models have been developed for a variety of medical disciplines. Chiropractic (Davenport, Cleland, & Kulig, 2009; Teyhen, Flynn, Childs, & Abraham, 2007), emergency medicine (Emparanza & Aginaga, 2001; Heyworth, 2003; Stiell et al., 1992), military medicine (Billings, 2004; Leisey, 2004; Mahieu, Witvrouw, Stevens, Van Tiggelen, & Roget, 2006; Rosin & Sinopoli, 1999; Springer, Arciero, Tenuta, & Taylor, 2000; Sutlive et al., 2008), physical therapy (Beneciuk et al., 2009; Childs & Cleland, 2006; Cleland et al., 2007; Hicks, Fritz, Delitto, & McGill, 2005; Iverson et al., 2008; Wainner et al., 2005), and orthopedics (Brenner, 2008; Flynn et al., 2002; Kuijpers et al., 2006; Leisey, 2004; Yuen, 2001), have all benefitted from their use. Some of the specific orthopedic conditions that clinical prediction rules have been implemented for include: ankle injuries (Emparanza & Aginaga, 2001; Heyworth, 2003; Rosin & Sinopoli, 1999; Stiell, 1996; Yuen, 2001), carpal tunnel syndrome (Wainner et al., 2005), cervical pain (Cleland et al., 2007), knee dysfunction (Iverson et al., 2008; Lesher et al., 2006), shoulder related conditions (Kuijpers et al., 2007; Kuijpers et al., 2006) , and low back pain (Childs et al., 2004; Flynn et al., 2002; Hicks et al., 2005; Iverson et al., 2008; Richardson et al., 2002).

Athletic training is lagging significantly in the area of clinical prediction modeling. A two-part series on how to create a clinical prediction rules was published in 2010 (Bruce &

Wilkerson, 2010a, 2010b). The first article involved the specifics of how to create a clinical prediction rule (Bruce & Wilkerson, 2010a). The second article in the series outlined a clinical prediction rule for overuse injuries in intercollegiate softball players (Bruce & Wilkerson, 2010b). The other two articles were both conducted at the University of Tennessee at Chattanooga with Dr. Gary Wilkerson as the lead author. The first study examined cardio-metabolic risks among intercollegiate football players (Wilkerson, Bullard, & Bartal, 2010). The second study was only a preliminary study, but looked at the ability to predict injuries to the core and lower extremity in intercollegiate football players (Wilkerson, Giles, & Seibel, 2012). Several other studies have been conducted at and have been presented as poster presentations, but have yet to be converted into manuscripts for publication in refereed journals (Burdette & Wilkerson, 2012; Cockrell & Bruce, 2008; Friess & Bruce, 2010; Henley, Bruce, & McDermott, 2012; Hess, Wilkerson, & Colston, 2012; Jones, Wilkerson, Colston, & Bruce, 2012; Karch, Wilkerson, & Bruce, 2012b; Michel, Colston, & Tanner, 2011; Reinecke & Wilkerson, 2012; Rigney & Bruce, 2010; Snider, MacLean IV, & Wilkerson, 2013; Stanley & Bruce, 2009; Tucker, Mullis, Wilkerson, & Bruce, 2013).

In the establishment of any prediction model the first two goals are to identify any and all potential predictor variables and to establish a clear operational definition of the dependent variable (Bruce, 2012; Bruce & Wilkerson, 2010a). The use of prediction modeling has utility for admission decisions for health care professions and for estimating success on a profession’s licensure or board exam since the outcome is dichotomous: (admitted to the program or not admitted to the program; passage of the exam or not passing the exam). Only two studies examined these two research questions and both studies were done in physical therapy, by Utzman, Riddle, and Jewell (2007a, 2007b).

For their first study, Utzman et al. (2007a) dichotomized their outcome variable, academic difficulty verses non-academic difficulty from the rating by program directors of the students in their program, both past and present. The predictor variables included uGPA, GREv and GREq, target year of graduation, age at time of admission, gender, race/ethnicity, and degree level. The researchers utilized a hierarchical logistic regression to control for confounding variables. They examined the Wald statistics and adjusted odds ratios to identify those variables that contributed to the prediction model significantly.

After running their logistic regression model, the authors then ran a receiver operator characteristic (ROC) curves to determine sensitivity and specificity. To develop cut-points the authors separated uGPA, GREv, and GREq into tertiles. They then recoded these variables and retested them against the previously identified predictor variables. If these tertile cut-points did not yield significant differences, then ROC curve analysis was used to identify cut-points (Utzman et al., 2007a).

In their study, tertile cut-points were used for the uGPA, while ROC curve analysis was used for GREv and GREq. To develop the final prediction model the β coefficients from the logistic regression was used to determine the strength of the variables that should be included. Their final analysis was to apply their model to a variety of physical therapy programs. The authors reported only the percentage of schools that their model fit rather than the sensitivity, specificity or odd ratios for each program (Utzman et al., 2007a).

In a second study, Utzman et al. (2007b) repeated their previous study, but this time to predict performance on the National Physical Therapy Examination (NPTE). They repeated the same procedures as they did in their previous study with the only difference being that they examined the odds ratios for failing the NPTE (Utzman et al., 2007a). “Odds ratios indicated that when controlling for other variables, the odds of failing the NPTE were increased 12% for each 0.10 decrease in uGPA. As GREv and GREq scores decreased by 10, odds of NPTE failure were increased by 6.6% and 3.5% respectively” (Utzman et al., 2007b, pp. 1185-1186). The authors concluded that their prediction model of uGPA, GREv and GREq was able to predict failure at least once on the NPTE. The GREv score was the strongest variable to predict failure on the NPTE. In their conclusions the researchers suggest that GREv and GREq are the strongest predictors for both failure on the NPTE and academic admission decisions (Utzman et al., 2007a).

There are a few problems with the two Utzman et al. (2007a, 2007b) studies. In both of the Utzman et al. studies, the authors use tertiles to determine cut-points, but how those cut-points were determined was not explained and appear to be arbitrary. For example, for the GPA, was it one-third of the entire 4.0 scale? Or was it one-third between 3.0 and 4.0? A better solution would have been to use the ROC curve analysis to determine the cut-points and to dichotomize high scores verses low scores (Bruce & Wilkerson, 2010a; Hosmer & Lemeshow, 2000). The dichotomized predictor variables would then be placed into a logistic regression for analysis of the best model. Instead Utzman et al. (2007a, 2007b), only used ROC curve analysis if tertile cut-points did not yield significant differences.

Another difficulty was the authors only used ROC curve analysis for determining sensitivity and specificity (Utzman et al., 2007a, 2007b). Dichotomizing their data would have permitted them to create a 2 X 2 cross-tabulation table and calculate the sensitivity, specificity, likelihood ratios, odds ratios and relative frequency for success (relative risk), from these data. They could have calculated a 2 X 2 cross-tabulation table for each of the predictor variables and obtained the information for each variable (Utzman et al., 2007a, 2007b).

Furthermore if the authors had dichotomized their data, they then could have determined who was above or below the established cut-points. (In this case, high scores and GPA would have been positive factors to gain admittance or to pass the NPTE.) After calculating the total number of positive factors an individual possesses, ROC curve analysis would be repeated to determine the optimum number of positive factors. With an optimum number of positive factors determined a prediction model could be developed and examined for sensitivity, specificity, and odds ratios (Bruce & Wilkerson, 2010a; Federation of State Boards of Physical Therapy, 2012; Wilkerson et al., 2010).

To develop the final prediction model Utzman et al. (2007a, 2007b) used β coefficients generated from the logistic regression. There were four better choices to determine which model fits best. The authors could have examined the chi-square statistic for significance, the Nagelkerke R^2 data for the amount of the variance the model accounted for, examined which step of the classification table provided the most accurate data to classifying subjects in their appropriate category: (true positives + true negatives / total), or the Exp(B) data to determine the odds ratios of each predictor variable and the interaction between each variable at each step of the prediction model.

Statistics Utilized for Prediction Modeling

A variety of statistical methods have been used to analyze the potential relationships between and among predictor variable to predict admission verses non-admission, or passing versus not passing a specific profession's certification/licensure examination. Multiple regression analysis was the most commonly used statistical technique implemented by a variety of authors and for various purposes (Balogun et al., 1986; Day, 1986; Hansen & Pozehl, 1995; Hayes et al., 1997; Julian, 2005; Kirchner & Holm, 1997; Kirchner et al., 1994; Levine et al., 1986; McGinnis, 1984; Meleca, 1995; Mitchell, 1990; Munro, 1985; Newton & Moore, 2007; Platt et al., 2001; Rhodes et al., 1994; Silver & Hodgson, 1997). Correlations were also used extensively. Several authors used the Pearson product-moment correlation coefficient (Hayes et al., 1997; Levine et al., 1986; Mitchell, 1990; Munro, 1985; Newton & Moore, 2007; Rhodes et al., 1994; Stricker & Huber, 1967), while the Spearman rho rank correlation coefficients were used in only one study (Morris & Farmer, 1999). A Pearson's Chi-squared tests was used in one study (Hickman, 2010)

Both the Draper (1989) and the Erickson & Martin (2000) studies reported only the specific percentage related to the data they collected. Draper (1989) reported scores for each section of the BOC exam and for the Learning Style Inventory scores. Although he does report the level of significance ($p < 0.05$), he does not state what statistical procedure was used to determine those p-values (Draper, 1989). Erickson & Martin (2000) descriptive study reports the percentages and means of the survey data they collected. Since they were only describing what athletic training education program director's believed contributors to success on the BOC exam, no p-values were reported (Erickson & Martin, 2000).

Harrelson et al. (1997) utilized a multiple linear regression to identify the predictor variables in relation to pass in the BOC exam pass rate. Multiple discriminant analysis was used to determine what, if any, combination of "variables could predict success" on the BOC exam (p. 325). Discriminant analysis assumes that the "predictor variables are distributed as a multivariate normal distribution with equal covariance matrix" (Peng et al., 2002, pp. 9).

Turocy et al., (2000) used "standard descriptive statistics, nonparametric analysis, parametric linear regression," and Pearson product-moment correlational analysis to examine their data (Turocy et al., 2000, pp. 71). No relationships were found for clinical experience hours, types of clinical experiences, or demographic information to predict scores or pass-fail outcome on the BOC exam or on any of the parts of the exam (Turocy et al., 2000).

A two-way analysis of variance was used to look at the differences between the means of candidates' scores by route to eligibility and by candidates' gender in the study conducted by Middlemas et al. (2001). Additionally, chi square analysis was performed to examine whether a difference existed between the internship and curriculum routes to certification. Correlations among the predictor variables were used to examine for collinearity among the predictors. To "determine the ability to predict the outcome" on each section of the exam "from the predictor variables" multiple regression analysis was used (Middlemas et al., 2001, p. 137). "Stepwise linear regression analysis was used to examine the ability to predict the quantitative score on each section of the certification examination from GPA and number of hours of clinical education completed" (Middlemas et al., 2001, pp. 137). Logistic regression was used to predict whether a candidate will pass or fail the entire BOC exam. The predictor variables used were uGPA, and clinical hours. Both variables were statistically significant and accounted for 58% percent of the variance ($R^2 = 0.58$) (Middlemas et al., 2001).

Hickman's dissertation (2010) examined if three variables were related to passing the BOC exam. She used contingency tables to assist her to decide if a relationship between the variables and the passing the BOC exam existed. A "statistical significance was noted if Prob>ChiSq was less than 0.05" (Hickman, 2010, p. 35).

Athletic training education program (ATEP) characteristics were examined and included "total number of clinical experience hours, the year in which the student was assigned their first rotation, and the number of clinical rotations assigned that consisted of more than 50 total hours" (Hickman, 2010, p. 35). She states her chi-squared (χ^2) analysis found no relationship between ATEP characteristics and success on the BOC exam, but does not report the specific findings (Hickman, 2010).

Although Hickman (2010) found that "four of five students who worked both preseason and fall football passed on the first attempts, while three of nine students who worked preseason football alone passed on the first attempt" there was no statistical relationship found (Prob>ChiSq= 0.086).

Age and GPA were student demographics which were also examined by Hickman (2010). She reports no statistically significant findings, but attributes it to her small sample size ($n = 24$). She reports the Prob>ChiSq = 0.081.

In all three of the variable analyses, Hickman (2010) uses a multiple column by two row contingency table. In both cases, ROC curve analysis would have given her a cut-point in order to dichotomize her data from a specific point rather than just arbitrarily selecting the cut-points.

For the football season experience X passing the BOC exam on the first attempt she uses eight categories. This creates some problems. In the notes of her contingency table she states that "20% of cells have expected count less than 5, Chi-square suspect" (p. 37). If she decreased her number of categories from eight to two thus, dichotomizing her data, it may have strengthened her analysis.

A re-configuration of Hickman's (2010) data to fit into a 2 X 2 was done as follows: All students who had experience working pre-season football camp and the fall football season were compared to all other categories of football. There were five students who worked both pre-season football and the fall season who passed the BOC exam and three who did not pass the BOC exam. There were five students who passed the BOC exam who had experience with all other combinations of football experience, while eleven students who worked other combination of football (FB) failed the BOC exam. Table 2.5 shows a reconfigured 2 X 2 contingency table.

Table 2.5 Football Experience X Passing vs. Not Passing the BOC exam

	Passed BOC exam	Failed BOC exam
Pre-season & Fall FB season experience	5	3
All other combinations of FB experience	5	11
Total	10	14F

Sensitivity: 0.50 (90% CI: 0.27 – 0.73)	+LR: 2.33 (90% CI: 0.87 – 6.28)
Specificity: 0.79 (90% CI: 0.57 – 0.91)	-LR: 0.64 (90% CI: 0.36 – 1.12)
OR: 3.67 (90% CI: 0.82 – 16.32)	RFS: 2.0 (90% CI: 1.63 – 2.45)

(Hickman, 2010).

The positive findings of these reconfigured data demonstrates the RFS is a 2.0 greater probability of an individual passing the BOC exam for those individuals with pre-season and fall football season experience compared to those who have any other combination of football experience. Although the odds ratio indicates an individual who has worked pre-season football and fall football has 3.67 times greater odds to pass the BOC exam than someone who has any other combination of football experiences, there are two problems. Since the 90% confidence interval is 0.82 – 16.32, thus crossing "the null value of 1.0," . . . "it can be concluded that the observed association is not statistically significant" (Hosmer & Lemeshow, 2000, p. 340)

A second problem is the same number of students who experienced both pre-season football and the fall football season passed the BOC exam as those who had any combination of football experience. Therefore, according to this small sample size in this study it appears a student's football experience has no bearing on passing the BOC exam (Hickman, 2010).

For her analysis of the GPA, she used 4 categories, so again if she had dichotomized the data she would have had a stronger analysis. In Hickman's (2010) data she sets her categories for "Adj. GPA" (but she never explained what Adj. GPA was or how it was calculated or determined) from 2.8 – 3.1; 3.2 – 3.5; 3.6 – 3.9; 3.9 – 4.2. (Note also that she does not account for the hundredths of a point would be classified in her Adj. GPA.) Hickman reports 14 students did not pass the BOC exam that had Adj. GPAs between 2.8 and 3.5, while six students did pass the BOC exam with Adj. GPAs in these categories. No students failed the BOC exam with an Adj. GPA between 3.6 and 4.2, and four students did pass the exam with an Adj. GPA in these categories. (Because there are no students who failed the BOC exam with an Adj. GPA between 3.6 and 4.2, a value of 0.5 is added to that cell and for consistency to all of the other cells too.

Otherwise, the odds ratio will be either zero or infinity (Hosmer & Lemeshow, 2000).) The new 2 X 2 contingency table is displayed in Table 2.6.

Table 2.6 Adj. GPA X Passing versus Failing the BOC exam

	Passed BOC exam	Failed BOC exam
3.6 – 4.2	6.5	0.5
2.8 – 3.5	4.5	14.5
Total	11	15
Sensitivity: 0.591 (90% CI: 0.35 – 0.79)	+LR: 17.73 (90% CI: 1.74 – 181.1)	
Specificity: 0.967 (90% CI: 0.80 – 0.995)	-LR: 0.42 (90% CI: 0.23 – 0.77)	
OR: 41.89 (90% CI: 3.2 – 548.4)	RFS: 3.92 (90% CI: 3.2 – 4.81)	

(Hickman, 2010).

The interpretation of this new configuration of Hickman's (2010) data demonstrates moderate sensitivity and excellent specificity. However, the odds ratio says an individual with an Adj. GPA of 3.6 – 4.2 is 41.89 times more likely to pass the BOC exam than someone that has an Adj. GPA between 2.8 & 3.5. Because of the small sample size, (two cells have less than five subjects per cell) the 95% confidence interval is very large (3.2 – 548.4). The relative frequency for success (RFS) tells us there is a 3.92 greater probability of an individual passing the BOC exam for those individuals with an Adj. GPA between 3.6 & 4.2 compared to those with an Adj. GPA of 2.8 to 3.5. Again, because of the small sample size (n = 24) the CI were quite large indicating large fluctuations are possible in the data.

Pickard's (2003) dissertation examined the role of mentorship of athletic training students and the effect it might have on the BOC exam. He used a variety of statistical procedures to analyze his data. He concluded that mentoring does not have an effect on the outcome of the BOC exam. An additional finding was the BOC exam did not measure mentoring relationships (Pickard, 2003).

Starkey and Henderson's (1995) study analyzed the differences between students who took the former internship route to certification and those students who graduated from an accredited athletic training education curriculum. This comparative study examined the results for the 1992 and 1993 calendar years. They reported the percentages for each route to certification and performance on each of the three sections of the exam and for passing all three portions. Students from accredited curriculum programs passed all three sections of the exam at a higher percentage than their internship route counterpart. Of those students who passed all three sections of the exam, 32% of those who came from a curriculum program passed compared to only 24.1% of those who came from an internship route to certification.

Additionally, t-tests were conducted comparing the two groups on each of the three sections of the exam and on the written section by each of the athletic training domains. (For the other two sections of the exam, responses to questions often encompassed multiple domains.). Each of these analyses were statistically significant at the 0.0001 alpha level (Starkey & Henderson, 1995).

The results of the Starkey and Henderson (1995) study were the impetus for the elimination of the internship route to certification. The NATA Educational Task Force made 18 recommendations to the NATA Board of Directors in 1996. The requirement of all candidates being from an accredited athletic training education program was announced in 1997 and went into effect in January of 2004 (Craig, 2003; Delforge & Behnke, 1999; Weidner & Henning, 2002).

William and Hadfield (2003) surveyed 60% of all athletic training education program directors to identify what attributes possessed by athletic training education programs and whether or not these attributes are related to their student's success on the BOC exam. Regression analysis and a general linear model statistics were used to analyze the results of her survey results. The following variables were analyzed and found to have a positive effect on the passing rate for the first-time pass rate on the BOC exam:

- Emphasis on teaching the seven athletic training domains and the competencies within each of the domains
- Having separate clinical and academic responsibilities for faculty and staff
- Avoiding the hiring of faculty members with K-12 teaching experience (Williams & Hadfield, 2003)

There were four variables which William and Hadfield (2003) identified as not being statistically significant to passing the BOC exam the first-time. They include:

- Grade point average
- The athletic training curriculum being associated with an allied health school
- The format in which course examinations were performed
- The existence of a capstone course (Williams & Hadfield, 2003)

Hanse and Pozehl (1995) examined admission criteria to predict achievement in a graduate level nursing program. They used factor analysis to evaluate the results from the

"Graduate Performance Rating Survey" utilized in their study. The researchers used step-wise multiple regression to examine the association between criterion variables and factor variables (Hansen & Pozehl, 1995).

Hayes et al., (1997), employed a variety of statistical methods to analyze their data. These included correlations, standard descriptive statistics, independent chi-square tests, multiple regressions, and t-tests. To determine if differences existed between their two groups, (traditional verses non-traditional students) they used independent t-tests. Additionally, the authors examined the correlations between all of the variables. Multiple regressions were used to analyze all students, traditional students and non-traditional students each "to determine which variables predicted academic success in the physical therapy program as determined by PT GPA" (Hayes et al., 1997, pp. 13). One problem with Hayes et al.'s model is that they do not identify a cut-point for what was considered a successful PT GPA.

Sensitivity/Specificity

Sensitivity (Sn) and specificity (Sp) is easily calculated by using a 2 X 2 cross-tabulations table. The four cells of the 2 X 2 cross-tabulation table are true positive, true negative, false positive, and false negative. How accurately a test is able to obtain a positive test when the actual condition is present is known as a true positive. When a test is able to identify a negative test when the condition is not present is known as a true negative. If a person is identified as potentially having a condition, but in actuality does not have the condition in question this is known as a false positive. A false negative is when a person is identified as not having a condition, but in actuality the diagnosis is positive (Munro, 2005b; Rothman et al., 2008).

The percentage of accurately identifying the number of true positives is known as Sn. Specificity is the ability of a test to classify those individuals without the condition. A test or instrument that is highly sensitive will rarely identify someone as positive if they do not have the condition. Likewise, a test or instrument that is highly specific will rarely identify someone as negative if they have the condition (Portney & Watkins, 2000; Rothman et al., 2008; Vincent & Weir, 2012). This relationship is shown in Table 2.7.

Table 2.7 Sensitivity and Specificity

	Gold Standard Test Positive	**Gold Standard Test Negative**	**Total**
Predicted Positive	True Positives (a)	False Positives (b)	a + b
Predicted Negative	False Negative (c)	True Negatives (d)	c + d
Total	a + c	b + d	Total Percentage Correctly Identified (a +d)/a + b + c + d

$$\text{Sensitivity} = \frac{a}{a + c} \qquad \text{Specificity} = \frac{d}{b + d}$$

$$\text{Odds Ratio} = \frac{a/c}{b/d} = \frac{ad}{bc} \qquad \text{Relative Risk}^{1} = \frac{a/(a + b)}{c/(c + c)}$$

For the purpose of this study the phrase Relative Frequency for Success (RFS) was substituted for Relative Risk.

ROC Curve Analysis

Receiver operator characteristic (ROC) curve analysis was developed during World War II to assist radar and sonar officers to determine what signals were actual ships or planes versus other miscellaneous noise, known as signal-to-noise ratios (Portney & Watkins, 2000). The sensitivity and specificity of actual signals versus other noise was represented on the ROC curve. An example of an ROC curve can be seen in Figure 2.4.

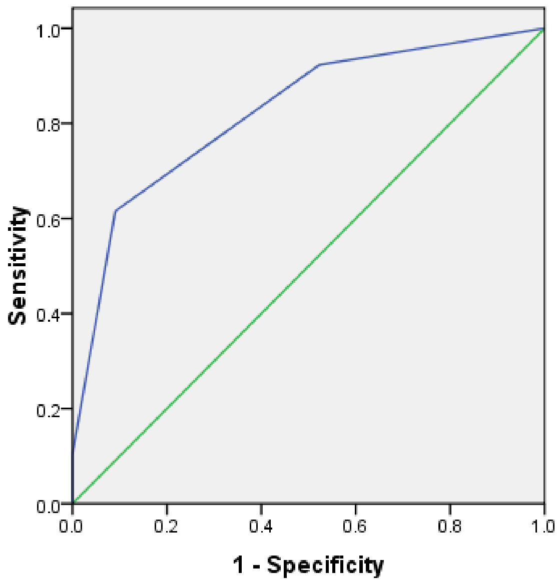

Figure 2.4 An example of a Receiver Operator Characteristic (ROC) curve showing the X and Y axis, for 1 – Specificity and Sensitivity respectively (Wilkerson, 2012).

An ROC curve can be created in the Statistical Package for the Social Sciences (IBM Corporation, 2011). On the ROC curve, the X-axis charts the 1-specificity (false positives), while the Y-axis charts the sensitivity (true positives). The X-Y intercepts represents the ratio of true positives and false positives. Unlike most graphing procedures where the further to the right on the graph data extends the more positive a test with an ROC curve a more positive result is seen toward the left. The point on the curve that is closest and approaches the upper left corner is usually selected as the cut-point. This means that your test has the highest possible sensitivity and a lowest possible 1 – specificity, (which calculates to an actual high specificity). The closer the curve is to the 45° reference line, the more likely the result is a 50/50 proposition (Hosmer & Lemeshow, 2000; Peng et al., 2002; Peng & So, 2002; Portney & Watkins, 2000).

The area between the 45° reference line and the curve is known as the area under the curve. This allows us to compare two or more criterion to determine which one might be better. A better test is represented by a higher area as it approaches 1.0. The area under the curve value is equal to the probability of correctly selecting the appropriate classification. Thus with an area under the curve of 0.852 would represent a correct choice of the criterion 85.2% of the time of randomly chosen subjects as seen in Figure 2.5 (Fawcett, 2006; Hosmer & Lemeshow, 2000; Portney & Watkins, 2000).

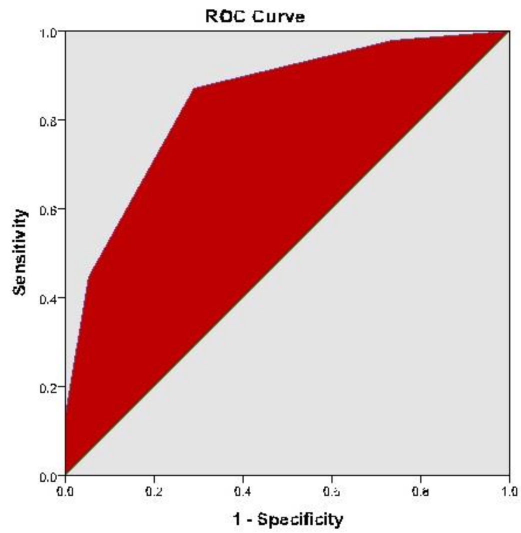

Figure 2.5 ROC curve with the Area Under the Curve (AUC) identified in red.

Youden's Index helps to provide an objective method of determining what the best point is on the ROC curve to provide the optimal value for variable discrimination (Ardern, Taylor, Feller, Whitehead, & Webster, 2013). In 1950, Dr. W. J. Youden saw a need to "reduce a table of data, into one figure that will adequately characterize (a) diagnostic test" (Youden, 1950, p. 32); hence, when looking at a series of 2 X 2 cross-tabulations tables with calculated sensitivity and specificity or an ROC curve with multiple of potential cut-points, Youden's Index is able to distinguish which point is the best cut-point. Youden's Index formula is:

$$J = \frac{ad - bc}{(a + b)(c + d)}$$

Where:

J = Youden's Index
a, b, c, and d are the cells of a 2 X 2 cross-tabulation table

(Youden, 1950)

An extension of Youden's Index occurred when Ruopp et al., (2008) reworked the formula so it became:

$$J = \max_c \left(Sn(c) + Sp(c) - 1\right)$$

Where:

J = Youden's Index
c = optimal cut-point for the Sn and Sp - 1
$\max_c$ = maximum cut-point on the ROC curve

(Ruopp et al., 2008)

Logistic Regression

Logistic regression has been used most effectively in educational and medical research (Hosmer & Lemeshow, 2000; Peng et al., 2002). It is based on the concept of "maximum likelihood,' meaning the procedure "will present the 'most likely' solution that demonstrates the best odds of achieving accurate prediction of group membership" (Portney & Watkins, 2000, p. 598) Logistic regression helps to determine the likelihood that a patient may fit into a high risk verses a low risk category. When confronted with a dichotomous outcome variable, logistic regression is the statistical procedure of choice. Because the outcome variable is categorical, it does not have normal or linear distributions, so neither multiple or linear regression can be used. During data entry, the condition is usually coded as zero (0) for a non-event, and one (1) for an event occurring (Field, 2009; Munro, 2005a; Portney & Watkins, 2000). The simplest result of a logistic regression is a 2 X 2 cross-tabulations table (Peng et al., 2002). Although continuous predictor variables provide a greater volume of information, they do not lend themselves to accurate placement into a 2 X 2 table (Table 2.3). By establishing cut-points through ROC curve analysis, dichotomizing the predictor variables as above or below the cut-point by recoding into

0 or 1, and classifying these data into the 2 X 2 table, it is easier for the researcher to establish the odds ratio (Wilkerson, 2011).

To determine the contribution of the predictor variables, the Wald statistic is used. The Wald statistic demonstrates whether or not the *b* coefficient for each specific predictor is significantly different from zero. When the coefficient is significantly different from zero, the assumption is the specific predictor is a significant contributor to the outcome (Field, 2009; Munro, 2005b)

Discriminant analysis or discriminant function analysis is another statistical measure that could be used for categorical outcome variables (Hosmer & Lemeshow, 2000; Portney & Watkins, 2000; Tabachnick & Fidell, 2007). It can be used with two or more groups that allows for classification of group membership. Discriminant analysis assumes that the predictor variables are normally distributed and their variances are equal across groups, while logistic regression makes no assumption regarding the distribution of the predictor variables. If mixtures of dichotomized and continuous variables are being used then logistic regression is obviously a better choice (Hosmer & Lemeshow, 2000; Portney & Watkins, 2000; Tabachnick & Fidell, 2007).

Predictor variable analysis can be accomplished several ways through most software packages. Stepwise selection is an efficient manner in which to screen a large number of variables to determine the best combination. To accomplish this either forward stepwise selection or backward stepwise selection can be used (Hosmer & Lemeshow, 2000; Portney & Watkins, 2000). The “selection or deletion of variables” is accomplished through “statistical algorithms" that examine each variable for their contribution to the model (Hosmer & Lemeshow, 2000, p. 116).

Forward stepwise selection adds each predictor variable based on the statistical significance the variable contributes to the model. If it significantly helps the model, the variable is retained. If the variable does not enhance the model, it is rejected (Hosmer & Lemeshow, 2000; Portney & Watkins, 2000). Forward selection is more likely than backward selection to exclude variables. This phenomenon is known as suppressor effects and occurs when a variable is contributing significantly to the model, but only if another predictor is held constant (Field, 2009).

Backward stepwise selection has been found to be most useful to the research efforts of the Graduate Athletic Training Program at the University of Tennessee at Chattanooga (Baldwin & Bruce, 2008; Bruce & Wilkerson, 2010b; Burdette & Wilkerson, 2012; Clark, Bruce, & Wilkerson, 2012; Cockrell & Bruce, 2008; Friess & Bruce, 2010; Henley et al., 2012; Hess, Wilkerson, & Colston, 2011; Jones et al., 2012; Karch, Wilkerson, & Bruce, 2012a; Michel et al., 2011; Morrison, Bruce, & Wilkerson, 2012; Rigney & Bruce, 2010; Snider et al., 2013; Tucker et al., 2013). This method allows the researcher to begin with all of the predictor variables to be examined as part of the logistic regression. The computer software eliminates the variable contributing the least at each step in the process until there are either no variables remaining or all of the remaining variables significantly contribute to the model. Backward stepwise selection lessens the risk of making a Type II error. The elimination occurs in one of three ways: the use of the likelihood ratio, conditional statistic (a less forceful variation of the likelihood ratio), and the Wald statistic (Field, 2009). A comparison between the specific steps in the model to the end step is made. The predictor variable whose relative importance among variables as determined from the *p*-value is found to be least helpful to the model is eliminated. The *p*-value used is not the traditional hypothesis testing value, but rather an indicator of the importance of the variables in the equation at that particular step. Backward stepwise regression analysis is likely to produce more variables for the model than forward stepwise regression selection. For this reason a more intensive scrutiny of the variables should be done (Field, 2009; Hosmer & Lemeshow, 2000).

In addition to ROC curve analysis predictor variables can be screened through parametric procedures. Independent t-tests, Mann-Whitney U tests and chi-square tests should be used for continuous, nominal and ordinal variables respectively. Since the purpose is not to determine significance for the predictor variables, but to screen variables for their potential predictive value, an alpha level as high as 0.15 or 0.20 can be set (Bruce & Wilkerson, 2010a; Kuijpers et al., 2006; Teyhen et al., 2007).

A common question is how many predictor variables are appropriate? Logistic regression is appropriate when five or more potential independent (predictor) variables have been selected or thought to be of value. Using logistic regression will allow the researcher to select the most appropriate variables (Childs & Cleland, 2006). It has been suggested that 10-15 positive events or subjects categorized as a “1” occur for each predictor variable identified in the prediction equation. Therefore, if three predictor variables have been identified then there should be 30-45 subjects classified as a “1” in the outcome variable. This helps to prevent large effect sizes and large confidence intervals that often occur as a result of small sample sizes (Childs & Cleland, 2006; Wasson et al., 1985).

Odds Ratio

The odds ratio is an estimate of how likely an individual belongs to a group for the event occurring compared to belonging to the non-event group with the presence of specific predictor variables (Field, 2009; Hosmer & Lemeshow, 2000; Portney & Watkins, 2000; Warren, 1971). An odds ratio of less than 1.0 indicates a decreased likelihood that an event will occur. An odds ratio of greater than 1.0 indicates an increased likelihood that an event will occur (Field, 2009; Peng et al., 2002; Peng & So, 2002; Tabachnick & Fidell, 2007; Warren, 1971). An odds ratio of 1.0 indicates that the "event has an equal chance of happening or not happening" (Warren, 1971, p. 937). Mathematically the odds ratio is expressed as:

$$\text{Odd} = P(\text{event}) / P(\text{no event})$$

where P(event) is the probability of the event occurring and P(no event) is the probability of the event from not occurring. In SPSS, "Exp(B)" is the adjusted odds ratio for the predictor variables as shown in Figure 2.6 (Field, 2009; Munro, 2005b).

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	read	.098	.025	15.199	1	.000	1.103
	science	.066	.027	5.867	1	.015	1.068

Figure 2.6 An example of the chart produced by SPSS showing the "Exp(B)" statistic.

However, an easier method to calculate the odds ratio is from the 2 X 2 table (Table 2.3). The odds ratio can be calculated as follows (Table 2.8):

Table 2.8 Equation for Odds Ratio

$$\text{Odds Ratio} = \frac{a/c}{b/d} = \frac{ad}{bc}$$

(Portney & Watkins, 2000)

The odds ratio interpretation is based on the premise that the desired outcome variable should be coded “1”. Regarding the present study, subjects accepted into the GATP and those candidates passing the BOC exam on their first attempt were coded as “1”.

Relative Frequency for Success

The odds ratio has wide use in epidemiological research. The interpretation of the odds ratio “is based on the fact that in many instances it approximates a quality called relative risk” (Hosmer & Lemeshow, 2000, p. 50). Relative risk (RR), “indicates the likelihood that someone who has been exposed to a risk factor will develop the disease, as compared with one who has not been exposed” (Portney & Watkins, 2000, p. 333), and it is used prospectively. In experimental research, the sample population at risk is allocated to a treatment group compared to a control group. The study is then conducted and the outcome is then observed (Portney & Watkins, 2000). Since this study will not be examining risk factors, and no known research has been found to develop a prediction model for acceptance into a program or predicting success on a board examination using the methods being used in this study, two new terms need to be identified.

A positive factor is identified as occurring if a subject has a score on a predictor variable that is greater than the established cut-point as determined through ROC curve analysis. For the purposes of this paper, instead of RR the phrase "relative frequency for success" (RFS) will replace RR. The operational definition for RFS was the likelihood the candidate who has been classified to be accepted into a GATP was accepted into the program compared to the candidate who has not been so classified. Additionally RFS will indicate the likelihood a participate who has been classified as predicted to pass the BOC exam will pass his/her board exam compared with one who has not been so classified. In reporting the results, a subject who has been classified into one of the two success categories with the specified number of positive factors is "X" number of times greater than those with less than the specified number of positive factors (Hosmer & Lemeshow, 2000).

Reliability of Grade Point Average

Several studies have examined the reliability of grades and grade point averages (Bretz, 1989; Clark, 1964; Etaugh, Etaugh, & Hurd, 1972; Morris & Farmer, 1999; Saupe & Eimers, 2012; Warren, 1971; Werts, Linn, & Jöreskog, 1978). Grade point average has been used as either a predictor of success (Armstrong et al., 1998; Balogun et al., 1986; Burton & Wang, 2005; Hocking & Piepenbrock, 2010; Kirchner & Holm, 1997; Kirchner et al., 1994; Kuncel, Hezlett, & Ones, 2001; Middlemas et al., 2001; Morris & Farmer, 1999; Morrison & Morrison, 1995 ; Stricker & Huber, 1967; Williams et al., 1970; Willingham, 1972) or as a criterion for admission into several professional programs (Bretz, 1989; Kuncel, Crede', & Thomas, 2007; Morrison & Morrison, 1995 ; Newton & Moore, 2007; Silver & Hodgson, 1997). Arguments concerning how accurate and reliable grade point averages are in relationship to the students' abilities and capabilities can be made on both sides.

Reliability involves how consistent an instrument or individual is in obtaining similar results over time (Portney & Watkins, 2000). Psychometricians have defined reliability as "the ratio of true-score variance to the sum of true-score plus error variance" (Singleton Jr. & Smith, 1978, p. 39).

Grades by their very nature have an element of subjectivity to them (Bailey, 2002). It is difficult to use test-retest reliability for a student in a specific class, since he/she would have already taken the class and been exposed to the material previously, one can assume that the student would earn a better grade. Intra-rater reliability, the ability of an individual to accurately measure across multiple trials is what would be most appropriate in grading (Portney & Watkins, 2000). Hence, the instructor would give the same grade to each student who earned a similar number of criteria needed for the grade across many semesters or years of teaching a specific course. A potential problem with intra-rater reliability is bias on the part of the instructor. Any number of subject criteria can influence how a teacher may assign grades. It is virtually impossible to blind the teacher to whom they are grading or assessing (Portney & Watkins, 2000). It may be possible to use the test/re-test approach to reliability, but it would have to be over the course of two or more semesters or years provided the same information and material is covered and measured the same way from one semester or year to the next (Saupe & Eimers, 2012).

Inter-rater reliability involves the ability of two or more evaluators to give the same grade to the same group of students (Portney & Watkins, 2000). The problems in providing like grades for the same course being taught by different instructors are numerous. The same deliverables and assessments could be required for the course, but how those elements are assessed or the emphasis placed on certain information over other information may vary from instructor to instructor. Basically, it comes down to consistently doing things from one semester or year to the next with as little change as possible (Warren, 1971).

Another issue when examining grades is the variance across institutions. For example, how students are graded at an Ivy League school is likely to be different than at a land grant university. Princeton University has a policy that only 35% of students in general studies courses and 55% of students in junior/senior level course are to receive an “A” (The Faculty Committee on Grading, 2005).

When examining reliability in grading, some basic approaches such as interclass correlation coefficient (ICC), split-half reliability, Cronbach’s alpha and Spearman-Brown statistics have been used. The analysis of variance (ANOVA), Cronbach’s alpha, and split-half reliability are all measures of internal consistency, that is, they measure the degree to which a test measures the same attribute or characteristic or combinations of multiple components of them (Portney & Watkins, 2000; Saupe & Eimers, 2012). “Because GPA is considered an indicator of overall academic achievement” . . . “internal consistency method(s)” are considered appropriate (Saupe & Eimers, 2012, p. 6). How to achieve this consistency is outside the scope of this study.

Another issue is related to the number of assignments that are used for grading. Reliability can remain high if the number of evaluations throughout the grading period is kept to a minimum. But with each increase in the number of assessment opportunities, the reliability of the grading decreases. This is especially true when more subjective assessments such as writing, presentations, and essay exams are used in the assessment of the students (Warren, 1971).

Finally, the problem of grade inflation and its impact upon grading cannot be ignored. All of the reasons for why grade inflation occurs is also beyond the scope of this paper, but it has been theorized that with grade inflation comes a decrease in the reliability of the grades awarded (Rojstaczer, 2009). As Figure 2.7 from Rojstaczer (2009) demonstrates there is a national trend of increasing GPAs, and with the advent of the plus-minus system for grading, an increase in grading categories, there has been an increase in the reliability coefficients. Figure 2.7 demonstrates:

> . . . the average undergraduate GPAs for American colleges and universities from 1991-2006 based on data from: Alabama, Appalachian State, Auburn, Brown, Bucknell, Carleton, Central Florida, Central Michigan, Centre, Colorado, Colorado State, Columbia, Cornell, CSU-Fullerton, CSU-Sacramento, CSU-San Bernardino, Dartmouth, Duke, Elon, Florida, Furman, Georgia Tech, Georgetown, Georgia, Hampden-Sydney, Harvard, Harvey Mudd, Hope, Houston, Indiana, Kansas, Kent State, Kenyon, Knox, Messiah, Michigan, Middlebury, Nebraska-Kearney, North Carolina State, North Carolina-Asheville, North Carolina-Chapel Hill, North Carolina-Greensboro, Northern Iowa, Northern Michigan, Ohio State, Penn State, Pomona, Princeton, Purdue, Roanoke, Rutgers, Southern Illinois, Texas, Texas A&M, Texas State, UC-Berkeley, UC-Irvine, UCLA, UC-Santa Barbara, Utah, UW-Oshkosh, Virginia, Washington State, Washington-Seattle, Western Washington, Wheaton (IL), William & Mary, Winthrop, Wisconsin-La Crosse, and Wisconsin-Madison. Note that inclusion in the average does not imply that an institution has significant inflation. Data on the GPAs for each institution can be found at the bottom of this web page. Institutions comprising this average were chosen strictly because they have either published their data or have sent their data to the author on GPA trends over the last 11-16 years. (Rojstaczer, 2009, para. 1)

In using grades for prediction modeling it is recommended that correlations and regression analysis be the statistics of choice (Rojstaczer, 2009).

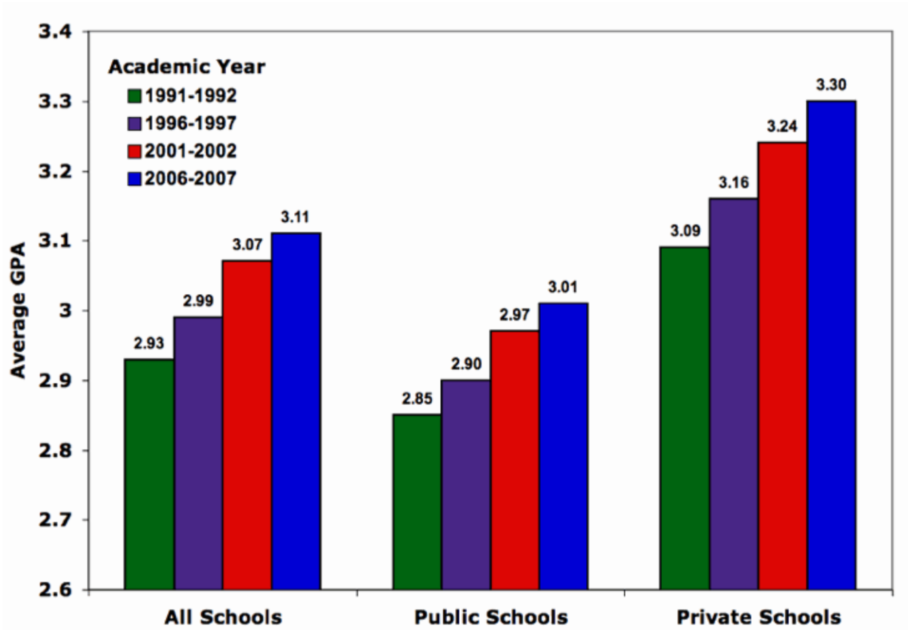

Figure 2.7 Recent GPA Trends Nationwide

(Rojstaczer, 2009 (used with permission))

Validity of Grade Point Average

Although according to the literature uGPA may not be reliable, several studies have shown uGPA to be a valid predictor (Etaugh et al., 1972; Kuncel et al., 2007; Kuncel et al., 2001; Morris & Farmer, 1999; Salvatori, 2001; Silver & Hodgson, 1997). Kuncel (2007; 2001) was the lead author for two meta-analyses. Both studies had a large number of student records to assess the predictive validity of not only uGPA, but the GRE (Kuncel et al., 2001) and the Graduate Management Admission Test (GMAT) (Kuncel et al., 2007).

Kuncel, Hezlett & Ones (2001) examined 82,659 student records and found that uGPA was a valid predictor of gGPA, especially when used in combination with the GRE. In the 2007 study Kuncel, Credè, & Ones examined the predictive validity of the GMAT and uGPA. They examined 64,583 students and determined GMAT was a superior predictor to uGPA separately and when combined the GMAT and uGPA were especially valid predictors. In a third meta-analysis in which Kuncel was a co-author (Grossbach & Kuncel, 2011) a total of 7,159 student records were examined and they determined that uGPA was a valid predictor for nursing students.

Overall, uGPA was found to be a valid predictor for several allied medical professions. These included: physical therapy (Burton & Wang, 2005; Day, 1986; Hayes et al., 1997; Kirchner et al., 1994; Levine et al., 1986; Shiyko & Pappas, 2009; Zipp et al., 2010); medical school (Cohen-Schotanus et al., 2006; Ferguson et al., 2002; Hamdy et al., 2006; Kreiter & Kreiter, 2007; Meleca, 1995; Salvatori, 2001; Silver & Hodgson, 1997); occupational therapy (Feldman, 2007; Kirchner & Holm, 1997); and nursing (Grossbach & Kuncel, 2011; Hansen & Pozehl, 1995)

Graduate Record Examination

The Graduate Record Examination (GRE) is a commonly used, standardized exam that has several purposes including admission decisions, preparedness for licensure or certification, course placement, employment decisions, and accountability for educational systems. It is most commonly used to assess a candidate's preparedness for graduate level work (Educational Testing Services, 2011b). The primary purpose of standardized testing is the ability they have to provide uniformity from one test group to another over the course of a similar time period and over a matter of years (Perdew, 2001; Risberg, 2010). This is supposed to be a measure of student achievement in their academic development; however, it fails to accomplish this objective. Standardized testing tends to say more about one's socioeconomic status than about the student's academic abilities (Wolk, 2009). A second purpose of standardized testing was the ability to compare large groups of students to make an accurate comparison between states to help determine which ones are having success and to assist teachers in where to direct their efforts to be most helpful in their teaching & educational strategies (Darling-Hammond & Rustique-Forrester, 2005 ; Hunsecker, 2007; Risberg, 2010).Standardized testing also provides the ability to compare students and applicants across different areas of the country or world. Students have different teachers and have different curriculums. Standardized testing is the only objective number provided for schools, colleges/universities, accrediting agencies, and other who may desire to study the results and make comparison across large populations. Grade point average provides a numerical assessment of a student's achievement and appears on all applications (Testing is Easy, n.d.). However, as we have stated earlier there are problems with the reliability of GPA due to variance across instructors and curriculums (Bretz, 1989; Etaugh et al., 1972; Morris & Farmer, 1999).

The Educational Testing Services (ETS) changed the manner in which the GRE was scored in 2011. The ETS provided concordance tables so GREs taken prior to 2011 could be compared to exams taken since 2011. Not only did the concordance tables allow old scores to be converted to new scores or vice versa, but also ETS provided percentile ranks of the scores (Educational Testing Services, 2011a). These percentiles were then modified slightly through April 2013. These revised scores and percentile ranks were used for the prediction models; thus, all candidates' scores, for all three sections, regardless of when they applied to the GATP, were converted to the most recent available scores and ranks (Educational Testing Services, 2013a, 2013b).

The GATP requires the General Test of the GRE which includes three sections: verbal reasoning (GREv), quantitative reasoning (GREq), and analytical writing (GREwr). Educational Testing Services (ETS), the organization responsible for administering the GRE revised the general test in August 2011. As with the previous version of the GRE General Test, ETS states that "the revised test measures the verbal reasoning, quantitative reasoning, critical thinking and analytical writing skills required for success in graduate and business school" (Educational Testing Services, 2011a, p. 4). The exam was revised in how the test was scored. Previously the general test was scored in 10 point increments from 200-800 points for each section. The revised GRE exam is currently scored from 130-170 points in one-point increments (Educational Testing Services, 2011a).

Reliability estimates for individual scores on the GRE revised General Test sections are as follows: verbal reasoning = 0.93; quantitative reasoning = 0.94; and analytical writing = 0.79. The standard error of measurements are 2.2, 2.0 and 0.4 for each section respectively (Educational Testing Services, 2011a). Data used to determine the percentile ranks were gathered from July 1, 2007 and June 30, 2011 and totaled more than 1.5 million examinees. According to ETS, "(The) percentile ranks are based on the concordance relationships between the prior 200-800 score scale to scores on the new 130-170 score scale. They are being used to provide stable and comparable interpretative information for scores on both scales" (Educational Testing Services, 2013b, p. 1).

Validity of the GRE has been established by several sources. Burton and Wang (2005) examined 21 graduate departments across seven different institutions. They established the use of the GRE with uGPA to determine ratings by faculty members, the student's first-year gGPA and the final overall gGPA. Kuncel et al. (2007; 2001; 2010) has conducted several studies regarding the GRE. A 2001 meta-analysis examined the ability of the GRE and uGPA to predict first year gGPA, faculty ratings, degree attainment, and scholarly productivity (Kuncel et al., 2001). Kuncel and Hezlett (2007) examined the ability of the GRE to predict success of graduate students. They also used the GRE to predict success on several standardized tests across several medical professions. They concluded that all standardized exams were able to predict success on the student's licensing exam, faculty ratings, research productivity, completion of their degree, their gGPA and first-year gGPA (Kuncel & Hezlett, 2007). A third meta-analysis conducted by Kuncel et al. (2010) studied the ability of the GRE to predict first-year gGPA, overall gGPA and faculty ratings in both master's degree programs and doctoral programs. The authors examined over 100 studies that included a combination of 1000 students and found the GRE to be very predictive of the predictor variables (Kuncel et al., 2010).

Chapter III

METHODS

Introduction

This study has two interrelated purposes, both of which pertain to the process of admitting students to a graduate professional program. The first component of this study involves the development of a prediction model to identify factors associated with eligibility and first-attempt success on the Board of Certification (BOC) examination for students who have completed a professional (entry-level) graduate athletic training program (GATP). The second component utilizes the results of the first analysis to identify program applicant characteristics that are most likely to predict both academic success within the graduate professional program and subsequent success on the BOC exam. In order to examine these two purposes Bayesian philosophy was used. Receiver operating characteristic (ROC) analysis was utilized to establish cut-points for each predictor variable, and logistic regression was used to assist in identification of the strongest combination of variables. Finally, a 2 X 2 cross-tabulations table was calculuated to determine the sensitivity, specificity, odds ratio and relative frequency for success.

Subjects

A cohort study design was used. The cohort consisted of students admitted to the GATP 2004 through 2012. The following information from a student's application folder was used in this study: all transcripts of undergraduate institutions attended to calculate uGPA and to determine whether or not the subject took advanced course work related to athletic training, the hard sciences and math courses, the type and number of such courses taken, GRE report showing the percentile ranks of the GREv, GREq, and GREwr scores, in-state versus out-of-state residency, and their degree granting institution. Based on the subject's degree granting institution, the Carnegie Classification was used to identify each school's academic and/or research classification. From each school's common data set, we determined the Academic Profile of Undergraduate Institutions (APUI) (Common Data Set Initiative, 2012; The Carnegie Foundation for the Advancement of Teaching, 2010).

The GATP prioritizes a minimum uGPA of 3.0 for further consideration for admission to the GATP. The GATP faculty has historically found students with an uGPA of less than 3.0 to have struggled more than students with an uGPA of greater than 3.0. A few exceptions have occurred for students who have undergraduate degrees from universities and colleges known for their high academic standards.

Descriptive statistics for the cohort are reported. Approximately 910 prospective students over 10 years have applied to the GATP, and 360 (or 40%) of these candidates had complete data sets. Of the original 910 applicants, 180 (20%) remained after eliminating students with an uGPA of less than 3.0. The cohort was further reduced based on those who were or were not offered a position in the GATP, which equaled roughly 130 students. These included students who have come into the GATP and for a variety of reasons left the program either voluntarily or have been academically disqualified from the program. The final cohort involved all students who entered the GATP stayed for at least the first-year, and those who dropped out and those who completed the GATP curriculum, and those who sat for the BOC exam from 2005-2013.

students, , including make up the cohort used for this study. Those students who entered the GATP but left the program regardless of the reason left after the first year are considered as part of the “fail” group related to passing or failing the BOC exam.

The study was submitted to the IRB committee for review and was approved. Application data were secured from the GATP which include uGPA, all GRE scores, degree granting undergraduate institutions, and applicants’ state of residence. To maintain anonymity, student identification numbers were assigned by the university’s Graduate School, which included a three letter, three digit code (i.e, abc123). For the initial coding, these identification numbers were used when provided, but in any cases when an identification number was not assigned a random identification number of the same style was be assigned. Once all data have been coded, all personal identifying information was stripped and discarded. In subsequent reports, data are reported only in the aggregate. All data were kept on a secure computer accessible only by the investigators. Student’s gGPA information at the end of the first year in GATP was gleaned from the university’s data base accessible to all faculty members.

Data Collection

Data were collected from candidates’ application folders. The following data were collected from each applicant’s folder: uGPA, GREv, GREq, GREwr scores, home state of residence, and degree granting undergraduate institution. Educational Testing Services (ETS) provides percentile rank scores for the raw GRE score data. Percentile rank scores are being used due to a change in the scoring system that was implemented in the August of 2011 by ETS.

Percentile rank scores are standardized across both scoring systems (Educational Testing Services, 2013a, 2013b).

The UTC Psychology Department uses a "Formula Score" to aid in the decision process for the selection of students to their graduate program. This Formula Score was created by a faculty member, Dr. Michael Biderman (Biderman, 2013). For the purposes of this study, this formula score will be referred to as Biderman's Formula Score. According to the UTC Psychology page, a score of 480 is considered average. The page informs the reader that scores "below 430 are less likely to be admitted than those with scores closer to 480." The information continues, "A student with a formula score above 480 will have a higher probability of being admitted" (Biderman, 2013, The formula score, para. 4). No specific statistics are provided to indicate how likely a candidate is to be accepted or not accepted into their program.

Biderman's Formula Scores were calculated from these data as follows: Biderman's Formula Score = (uGPA x 100) + GREv PR + GREq PR + GREwr PR (Biderman, 2013). Means and standard deviations were determined for the cohort for all of the continuous and multi-level discrete predictor variables. A college's or university's status (private versus public), were coded as ones ("1") and zeros ("0") respectively. An institution's basic academic rating as determined from *The Carnegie Classification of Institutions of Higher Education*™ (The Carnegie Foundation for the Advancement of Teaching, 2010) was dichotomized in a variety of ways to isolate a single classification versus all other classifications to identify its strength as a predictor. The classification of interest was always coded as a one ("1"), while all others were coded as zero ("0"). Graduate GPA (gGPA) at the conclusion of a student's first year were obtained from the university's data base accessible to all faculty members. The results of students first-attempt taking BOC exam were provided by the GATP Director.

An analysis was conducted to determine colleges/universities with high academic standards versus those with less than high academic standards. This became known as the Academic Profile of Undergraduate Institutions (APUI). Those students who were offered a position, accepted the position, remained in the GATP for at least the first year, and either dropped out or were academically disqualified along with those who completed the GATP were part of the group of subjects used for the prediction model. Subjects who pass the BOC exam on the first-attempt were coded as a one ("1") while those who failed on their initial attempt taking the BOC exam or either drop out before completing the GATP curriculum, or were academically disqualified were classified as unsuccessful and were coded as zero ("0").

Determination of Academic Profile of Undergraduate Institutions

In order to quantify the Academic Profile of Undergraduate Institutions (APUI) from which students received their undergraduate degrees, each college or university in which a student graduated, and who accepted a position in the GATP, and completed at least the first-year in the GATP was included in this analysis. If the student had a gGPA at the end of their first year in the GATP of ≥ 3.45 they received a code of "1", while students with a gGPA of < 3.45 were coded with a "0".

The Google search engine was used to search for each school's web site. On the initial results page from Google, a brief profile of the school was provided and within this profile was each school's acceptance rate. This was used as one of the independent variables to determine the APUI.

Within each school's web site, a search for the mean or median ACT and/or SAT score was sought. In most cases for institutions that participated in the Common Data Set Initiative the information was found by doing a search for the "Common Data Set" (Common Data Set Initiative, 2012). In cases where multiple years of reports were available the most current year's report available was used. There were cases in which institutions reported only the ACT or SAT the data, but not both. In these cases, only the reported standardized exam data were recorded, but if an institution reported both set of exam scores, both were recorded. The data were provided in one of three ways: via a range of 25th to 75th percentile, as the median of all test scores, or as the mean of all test scores. In those cases where schools did not participate in the Common Data Set Initiative, their ACT and SAT scores may have been published in other locations on the institution's web site such as through the Admission's Office or through the "Quick Facts" or "Fast Facts" page. There were several cases where this information could not be found on the school's web site; thus, a search was made on the About.com College Admission web site (About.com, 2013). The "mean/median" was achieved by either using the reported mean of each institution's ACT/SAT score or the middle score of the reported 25th and 75th percentile scores. Once all Institutions' ACT and SAT information was located and entered onto the spreadsheet, the data were downloaded into IBM SPSS 20 (IBM Corporation, 2011) for statistical analysis. The mean, median and standard deviation of the Institutions' ACT and Institutions' SAT mean/median scores were determined along with the calculated 75th percentile and 80th percentiles.

Receiver operator characteristic analyses were done on each of the potential individual predictors to determine the best balance between sensitivity (Sn) and specificity (Sp) to establish the optimum cut-points for the purpose of dichotomizing each predictor. Based on the established cut-points subjects received a one (“1”) if they earned a score of greater than or equal to the cut-point and a zero (“0”) if they earned a scored below the cut-point. Cross-tabulation calculations were performed for the coded values of the various cut-points of each predictor. The cross-tabulation calculations generated Sn, Sp, odds ratio (OR), the relative frequency for success (RFS), and the *p*-value for Fisher’s Exact Test (one-sided). These data were used to determine the Academic Profile of Undergraduate Institutions (APUI).

Statistical Analysis

Once all data have been collected and entered onto a spreadsheet it was cleaned so as to eliminate those students who do not meet the inclusion criteria. For passing the BOC exam on the first attempt prediction model, the inclusion criteria were those students who were offered and accepted positions in the GATP, and completed at least the first-year of study in the GATP. If after the first-year in the GATP, a student dropped out of the GATP or were academically disqualified they were classified as failures for the first-attempt on the BOC exam. To predict success in the GATP, candidates with a completed file, who received an offer to be a part of the GATP, accepted the offer, and remained in the GATP for at least the first-year were included in the sample. Means and standard deviations for demographic data and the predictor variables were reported for all candidates. Data analysis for both prediction models was achieved through SPSS Statistical Package for the Social Sciences (IBM Corporation, 2011).

The literature suggests continuous predictor variables be entered into a logistic regression (Flynn et al., 2002; McLean Jr., 1969; Melendez, Bruce, & Wilkerson, 2010; Wilkerson et al., 2010). Hosmer and Lemeshow (2000) state entering continuous predictor variables is acceptable; however, if continuous predictor variables are used then "a meaningful change must be defined" (p. 64). The GATP faculty have found a more efficient way to handle the entry of the predictor variables into the logistic regression by dichotomizing each predictor (Burdette & Wilkerson, 2012; Cockrell & Bruce, 2008; Friess & Bruce, 2010; Henley et al., 2012; Hess et al., 2012; Jones et al., 2012; Karch et al., 2012b; Michel et al., 2011; Reinecke & Wilkerson, 2012; Rigney & Bruce, 2010; Snider et al., 2013; Stanley & Bruce, 2009; Tucker et al., 2013).

To accomplish this goal of dichotomizing each predictor variable, ROC analysis for all multi-level discrete and continuous variables was conducted. "Optimum cut-points" for dichotomizing these data were determined by calculating Youden's Index, the difference of the sensitive minus the 1-specificity figures provided by the "Coordinate on the Curve" table from SPSS (Böhning, Böhning, & Holling, 2008). Youden's Index provides an objective measure for the optimum cut-point on the ROC curve which is the point closest to the upper left hand corner of the graph for ROC analysis (Ardern et al., 2013; Böhning et al., 2008).

To assess for multicollinearity, linear regressions were utilized to examine the relationship between the independent variables (Field, 2009; Mertler & Vannetta, 2005b). Potential independent variables to be used in the prediction model were initially examined as continuous or multi-level discrete variables. This was followed by the dichotomized version of the continuous and multi-level discrete variables based on their cut-points plus any originally dichotomized variables.

To evaluate for interaction effects of the predictor variables three methods were used:

1. Graphic representation of the interaction between two predictors
2. Combining predictors and examining through a 2 X 2 cross-tabulation table
3. Calculate the Mantel-Haenszel common OR and the Breslow-Day tests for homogeneity

Coding

Candidates who scored at or above the designated cut-point on the specific predictor variable received a code of one (1) and if they scored below the designated cut-point they received a code of zero (0). Coding in this way permitted the creation of 2 X 2 cross-tabulation tables. Dichotomizing the predictor variables is appropriate since the outcome variable is dichotomized (Hess et al., 2011; Keskula et al., 1995; Masters, 1974; Rojstaczer, 2009; Singleton Jr. & Smith, 1978; The Faculty Committee on Grading, 2005). The following predictor variables were dichotomized: institutional control, candidate’s residency, individual basic Carnegie classification categories, size and settings, specific athletic training courses, and advanced math and science courses. Table 3.1 summarizes the coding used for these variables.

Table 3.1 Coding of dichotomized independent variables

Offered a position = 1; Not offered a position = 0
Institutional control: Public = 1; Private = 0
Candidate's residency: In-state = 1; Out-of-state = 0
Basic Carnegie classification categories
Bachelors only = 1; Others = 0
Bachelors and Masters = 1; Others = 1
Doctorate/Research = 1; Others = 0
Research Intensive = 1; Others = 0
Size & setting: Large (10,000+ undergraduates) = 1; Others = 0
Size & setting: Medium (3000-9999 undergraduates) = 1; Others = 0
Size & setting: Small (<1000-2999 undergraduates) = 1; Others = 0
Did the candidate take:
Advanced coursework: 1 = Yes; 0 = No
Athletic training coursework: 1 = Yes; 0 = No
Basic athletic training or Care & Pre courses: 1 = Yes; 0 = No
Advanced athletic training courses: 1 = Yes; 0 = No
Biomechanics: 1 = Yes; 0 = No
Advanced Sciences & Math Coursework: 1 = Yes; 0 = No
Any advanced biology: 1 = Yes; 0 = No
Any advanced chemistry: 1 = Yes; 0 = No
Calculus: 1 = Yes; 0 = No
Pathophysiology: 1 = Yes; 0 = No
Physics: 1 = Yes; 0 = No

Next, 2 X 2 cross-tabulation, univariable analysis was conducted to examine each predictor variable for its potential value for the multivariable analysis. Those predictors with an OR of greater than or equal to 2.0 (Hosmer & Lemeshow, 2000; Portney & Watkins, 2000) or a *p*-value for the Fisher's Exact Test (one-sided) ≤ 0.20 were considered for the multivariable analysis (Bruce & Wilkerson, 2010a; Kuijpers et al., 2006; Teyhen et al., 2007). Each of the individual predictors was then entered into a logistic regression to assess for the strongest set of predictors. The remaining set of individual predictors from the univariable logistic regression was then summed for each subject to determine the number of positive predictors he or she possessed. This was known as the total number of positive factors for that individual. Another ROC analysis was conducted to determine the optimum number of positive factors for the prediction model. Based on this cut-point each subject received a "1" if the number of positive factors each possessed was equal to or greater than the cut-point value. If a subject has fewer positive factors than the cut-point value, he or she was given a "0". Finally, a 2 X 2 cross-tabulation table was created along with it associated statistics.

Multicollinearity

A series of linear regressions was performed on the multi-level discrete and continuous variables to examine for multicollinearity, which occurs when predictor variables highly correlate to each other ($r \geq 0.80$-0.90) (Field, 2009; Mertler & Vannetta, 2005a; Portney & Watkins, 2000). Two statistical results are produced by SPSS through its collinearity diagnostics function: variance inflation factor (VIF) and tolerance. The VIF signifies the presence of a strong linear relationship between predictor variables. Both Field (2009) and Mertler & Vannetta (2005b) state although there is no hard evidence of a specific VIF value that should cause concern, they do agree a value of ten or greater indicates collinearity. Tolerance is the inverse of the VIF (1/VIF), thus values of < 0.1 should be a matter of concern (Field, 2009; Mertler & Vannetta, 2005a).

The multi-level discrete and continuous variables included in the multicollinearity analysis were the percentile ranks of the GREv, GREq, and GREwr scores, Revised GRE – Composite score, uGPA, the number of advanced math and science courses, total number of advanced math, science and athletic training courses, APUI score, and Biderman's Formula Score. The process was then repeated on the dichotomized version of the multi-level discrete and continuous predictor variables using the established cut-points along with the other dichotomized variables. These included the following: whether or not the student took physics as an undergraduate, whether or not the student took calculus as an undergraduate, and whether or not the student's undergraduate institution was classified as a research intensive through the Carnegie Classification system.

A logistic regression was performed to determine the strongest set of predictors and to examine for the interaction effects. The adjusted odd ratio (Adj OR), ("Exp(B)" in the SPSS analysis), was used to further interpret the interaction between the various predictor variables upon the outcome variables. The advantage is the researcher can determine the strongest "predictor variables . . . associated with the outcome" (Laupacis et al., 1997, p. 491).

Receiver operating characteristic analysis with Youden's Index calculations was performed to identify the optimum number of predictor variables "that offers the most accurate prediction" model (Wilkerson et al., 2010, p. 69). With the determination of the optimum number of factors a 2 X 2 cross-tabulation table was used to calculate the requisite statistics. Based on the results of these data a prediction model was created for predicting BOC exam performance and success in a GATP.

Interaction Effects

Statistically there are two types of effects: the main effect and the interaction effect. The main effect is the result each individual predictor has on the outcome variable. Because with multivariable analysis more than one predictor can have an effect on the outcome variable, the concept of confounding can occur, meaning “the observed effect could (possibly be) caused by something else” (Verhagen & Van Mechelen, 2009, p. 37). Verhagen and van Mechelen (2009) offer an unofficial rule stating that “when the regression coefficient of interest changes with more than 10%, there is relevant confounding” (p. 37). Therefore, a final important step in the use of logistic regression for prediction modeling is to examine for interaction effects (Hosmer & Lemeshow, 2000).

Where multicollinearity examines overlap or the correlation between predictor variables, interaction effects examine how one variable acts upon on all other variables in the model (Field, 2009; Hosmer & Lemeshow, 2000; Portney & Watkins, 2000). The purpose of assessing for interaction effects “is to determine whether or not the odds ratios are constant, or homogeneous, over the strata” (Hosmer & Lemeshow, 2000, p. 79). The interaction occurs when the relationship between variables is linear, but the slopes of the lines differ. When represented graphically, if the lines of two variables do not intersect there is an absences of interaction effect between those two variables. But if the lines do bisect, then an interaction effect is present between the two variables (Hosmer & Lemeshow, 2000; Portney & Watkins, 2000).

To correct for variations in a logistic regression an examination of the adjusted odds ratio should be done. The adjusted odds ratio takes into consideration the effect of two or more predictor variables have on the outcome variable (Portney & Watkins, 2000). The adjusted odds ratio reveals how the odds ratio is altered to determine the impact each of the multiple predictor variables has on the outcome variable. In SPSS, the output lists the adjusted odds ratio as "Exp(B)". The adjusted odds ratio is "the change in odds resulting from a unit change in the predictor" (Field, 2009, p. 270).

According to the literature, examining the adjusted odds ratio is not enough to assess for interaction effects. There are three additional ways in which to scrutinize for interaction effects between variables. The first is to graphically represent the interaction to examine if the lines of the two predictor variables intersect (Portney & Watkins, 2000). A second method is to prepare a list of all possible interactions between any two variables from the final logistic regression model and assess through a 2 X 2 cross-tabulation table (Hosmer & Lemeshow, 2000). The final methods are to use the Mantel-Haenszel common OR estimator equation (Hosmer & Lemeshow, 2000; Portney & Watkins, 2000) and the Breslow-Day Test for homogeneity (Lai, Mink, & Pasta, n.d.; Prieto-Marañón, Aguerri, Galibert, & Attorresi, 2012). "The Mantel-Haenszel estimator is a weighted average of the stratum specific odds ratio" and is made up of "the observed cell frequencies in a 2 X 2 table" for each stratum (Hosmer & Lemeshow, 2000, p. 80), (Table 3.2)

Table 3.2 Mantel-Haenszel estimator equation

$$OR_{MH} = \frac{\sum a_i \text{ x } d_i / N}{\sum b_i \text{ x } c_i / N}$$

(Hosmer & Lemeshow, 2000)

The Breslow-Day test is used to assess for homogeneity of the stratum-specific odds ratios. For the test to be valid, each of the cells in a 2 X 2 table should have a count of greater than five, thus it takes large sample sizes for each of the stratum examined (Lai et al., n.d.; Prieto-Marañón et al., 2012), (Table 3.3).

Table 3.3 Breslow-Day test for homogeneity of the odds ratio

$$BD = \sum_{j=1}^{I} \frac{(a_j - A_j(OR_c))^2}{Var(a_j; OR_c)}$$

(Prieto-Marañón et al., 2012)

All three methods of assessment were done for the analysis of the prediction model for GATP applicant success.

CHAPTER IV

RESULTS

This study had two interrelated purposes, both of which pertained to the process of admitting students to a professional graduate athletic training program. The first component of this study involved the development of a prediction model to identify factors associated with eligibility and first-attempt success on the Board of Certification (BOC) examination for students who have enrolled in a professional (entry-level) graduate athletic training program (GATP). The second component utilized the results of the first analysis to identify program applicant characteristics that are most likely to predict both academic success in the graduate professional program and subsequent success on the BOC exam. This chapter presents the statistical testing and results.

Predicted BOC Exam Success as a Criterion for GATP Admission

According to Stephen Covey's *7 Habits for Highly Effective People* (2004), Habit 2 is that one should "begin with the end in mind." From this perspective, the culmination of a student's athletic training education is to become eligible to take the BOC exam and pass the exam on the first-attempt. A new accreditation standard of the Commission on Accreditation of Athletic Training Education (CAATE) states all programs must publish student outcome data on their web site home pages. This is to include the number of students graduating from the program who took the BOC exam, the percentage of students who have passed the exam on the first-attempt, and the number of students who ultimately passed the exam, regardless of the number of attempts. According to CAATE, programs that do not have a three-year aggregate first-time pass rate ≥ 70% are said to be “in non-compliance” (Commission on Accreditation of Athletic Training Education, 2013b, "Becoming an Athletic Trainer", 3rd question, 5th bullet point). Thus, passing the BOC exam on the first-attempt is the program outcome of primary importance.

Descriptive statistics for students who completed the first year in the GATP, and who subsequently took the BOC examination, are presented in Table 4.1.

Table 4.1 Descriptive statistics for students enrolled in the GATP

		gGPA at the end of the First-yr.	uGPA	GRE Composite	GREv	GREq	GREwr
N	**Valid**	119	119	115	115	115	106
	Missing	0	0	4	4	4	13
Mean (± sd)		3.59 (± 0.38)	3.27 (± 0.29)	293.40 (± 9.04)	147.97 (± 5.11)	145.43 (± 5.16)	3.887 (± 0.68)
Median		3.67	3.23	293.00	148.00	145.00	4.0

		GREv PR	GREq PR	GREwr PR	[a]Biderman's Formula Score
N	**Valid**	115	115	106	106
	Missing	4	4	13	13
Mean (± sd)		37.57 (± 19.14)	25.63 (± 17.13)	50.32 (± 24.93)	443.16 (± 63.97)
Median		36.00	22.00	54.00	441.500

Note. gGPA = Graduate Grade Point Average; uGPA = Undergraduate Grade Point Average; GRE = Graduate Record Exam; GREv = Verbal section of the Graduate Record Examination; GREq = Quantitative section of the Graduate Record Examination; GREwr = Analytical Writing section of the Graduate Record Examination; GREq PR = Percentile Rank of the Quantitative section of the Graduate Record Examination; GREv PR = Percentile Rank of the Verbal section of the Graduate Record Examination; GREwr PR = Percentile Rank of the Analytical Writing section of the Graduate Record Examination

[a]Biderman's Formula Score = (100 * uGPA) + GREv PR + GREq PR + GREwr

The first step of the process for development of a prediction model was the performance of univariable analyses for factors believed to forecast first-attempt success on the BOC examination. The most commonly accepted indicator of academic success is grade point average (GPA). A receiver operating characteristic (ROC) analysis was performed for graduate grade point average (gGPA) at the completion of the first year of the two-year graduate program, using success on the BOC exam on the first-attempt (Yes or No) as the dichotomized outcome variable (Note: The definition of "No" includes students who gained eligibility to take BOC exam, but failed the exam on their first-attempt; students who failed to attain eligibility either because they dropped out of the GATP after the first-year, or they were declared academically deficient). A total of 136 students took the BOC exam. A GATP student was classified as successful on the BOC exam if they passed on the first-attempt taking the exam (n = 90). Students who either failed the BOC exam on the first-attempt taking the exam (n = 24) or those who dropped out of the program after their first-year in the GATP (n = 5) were classified as not being successful on the BOC exam on their first-attempt. The result of this analysis is presented in Figure 4.1 and Table 4.2. A cut-point of gGPA $\geq$ 3.45 was found to provide the best balance of sensitivity and specificity for prediction of first-attempt success on the BOC examination.

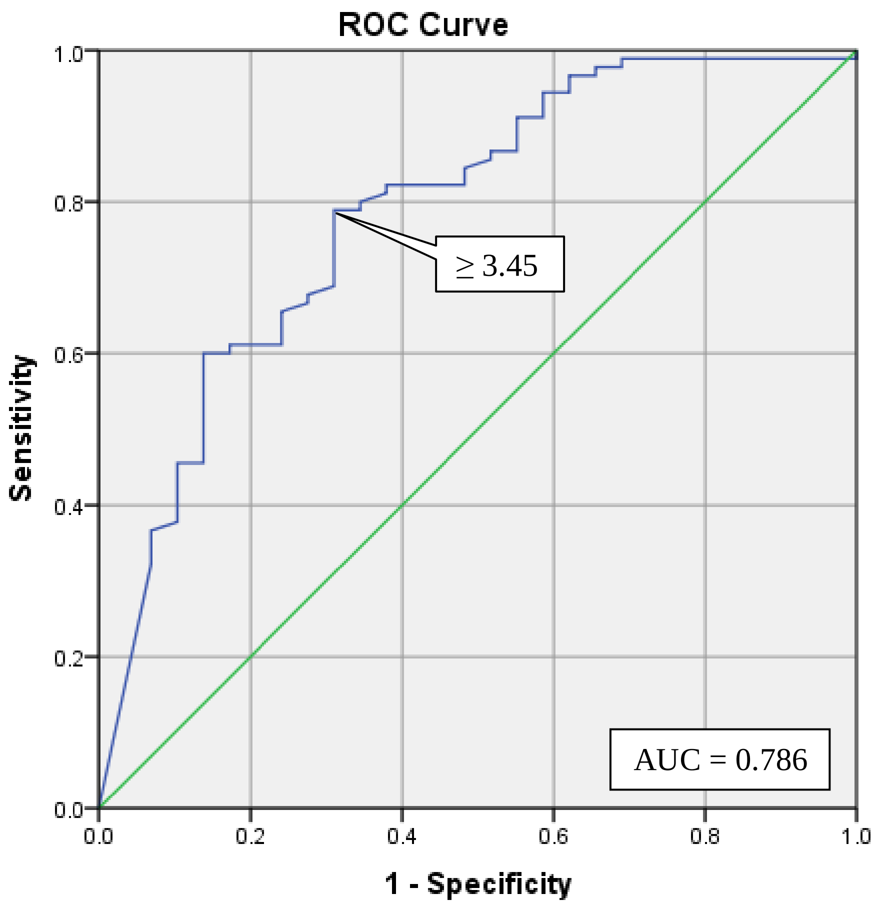

Figure 4.1 ROC curve with identification of the optimum cut-point for first-year gGPA as a predictor of first-attempt BOC exam success

Table 4.2 First-year gGPA for prediction of first-attempt pass – Yes or No, on the BOC exam

<table>
<tr><th rowspan="2"></th><th colspan="3">First-attempt Pass on the BOC exam</th></tr>
<tr><th colspan="2">Yes</th><th>No</th></tr>
<tr><td>First-year gGPA ≥ 3.45</td><td colspan="2">71</td><td>9</td></tr>
<tr><td>First-year gGPA < 3.45</td><td colspan="2">19</td><td>20</td></tr>
<tr><td colspan="4">Fisher’s Exact Test (one-sided) p < 0.001</td></tr>
<tr><td colspan="2">Sn = 0.79 (95% CI: 0.69 – 0.86)</td><td colspan="2">Sp = 0.69 (95% CI: 0.51 – 0.83)</td></tr>
<tr><td colspan="4">Youden’s Index = 0.479</td></tr>
<tr><td colspan="2">OR = 8.30 (95% CI: 3.26 – 21.16)</td><td colspan="2">RFS = 1.82 (95% CI: 1.49 – 2.23)</td></tr>
</table>

This analysis indicated that a student who had a gGPA ≥ 3.45 at the end of the first year had 8.30 times greater odds of passing the BOC exam on the first-attempt than the odds for someone who had a gGPA < 3.45 at the end of the first year. The relative frequency of GATP success indicates the probability of a student passing the BOC exam on the first-attempt with a gGPA ≥ 3.45 at the end of the first year is slightly less than twice the probability of a student with a gGPA < 3.45.

Several other variables were analyzed in an attempt to predict first-attempt success on the BOC exam. Receiver operating characteristic analysis was performed to determine the optimum cut-point for each possible predictor, along with 2 X 2 cross tabulation analysis to generate values for sensitivity (Sn), specificity (Sp), odds ratio (OR), the relative frequency for success (RFS), and the *p*-value for Fisher's Exact Test (one-sided). Each subject who had a score on a potential predictor variable greater than or equal to the cut-point was coded as a "1". If the student scored below the cut-point value, he or she was coded with a "0". Receiver operating characteristic (ROC) analysis results includes the area under the curve (AUC), Sn, and 1-Sp. Youden's Index is calculated from the Sn and 1-Sp values (Böhning et al., 2008; Ruopp et al., 2008). The 2 X 2 cross-tabulation analysis provides corresponding Sn and Sp values for the cut-point identified by the greatest value of Youden's index, as well as OR and RFS values. The univariable analyses for the potential predictors related to first-attempt pass – Yes or No, on the BOC exam are included in Appendix A, and the related information is summarized in Table 4.3 with the variables listed in the order of the OR magnitude.

Table 4.3 Summary of univariable results for potential predictor variables of first-attempt BOC exam success

Variable – First-attempt pass – Yes or No, on the BOC exam	**Cut-point**	**Sn**	**1 - Sp**	**Sp**	**Youden's Index**	**AUC**	**OR**	**RFS**	**Fisher's Exact Test (one-sided) *p*-value**
gGPA at end of the first year	3.45	0.79	0.31	0.69	0.480	0.786	8.30	1.82	0.001
GREq (PR)	143.5 (16.5)	0.72	0.31	0.69	0.411	0.758	5.76	1.53	0.001
GRE – Composite	290.5	0.70	0.31	0.69	0.389	0.736	5.17	1.48	0.001
Biderman's Formula Score	420.5	0.69	0.32	0.68	0.372	0.698	4.78	1.41	0.003
GREwr (PR)	3.25 (24.5)	0.89	0.64	0.36	0.257	0.587	4.76	1.59	0.007
GREv score (PR)	145.5 (26)	0.78	0.46	0.54	0.538	0.682	4.25	1.45	0.005
Number of advanced math, science or athletic training courses	3.5	0.62	0.34	0.66	0.273	0.640	3.07	1.32	0.017
Number of advanced math and science courses	2.5	0.51	0.30	0.70	0.196	0.586	2.27	1.21	0.087

Note. For further consideration a variable had to have an OR of ≥ 1.50 (Hosmer & Lemeshow, 2000) and a Fisher's Exact Test (one-sided) *p*-value of ≤ 0.20 (Bruce & Wilkerson, 2010a; Kuijpers et al., 2006; Teyhen et al., 2007)

Multicollinearity

A series of linear regression analyses were performed on the multi-level discrete, continuous, and dichotomous variables. These included: GREq, GRE – Composite score, Biderman's Formula Score, GREwr, GREv, total number of advanced science, math, and athletic training courses taken, and the number of advanced math and science courses taken. The multicollinearity analysis results for continuous and multi-level discrete variables, including tolerance and variance inflation factor (VIF) values are presented in Table 4.4.

Table 4.4 Multicollinearity analysis results

	Multicollinearity Statistics	
	Tolerance	**VIF**
gGPA at the end of the First-yr	0.563	1.775
Number of Adv Math & Science Courses	0.188	5.311
Total Number of Adv Courses (AT + Adv Math & Science)	0.187	5.348
GREv	0.395	2.532
GREq	0.385	2.596
GREwr	0.463	2.158
Biderman Formula Score	0.170	5.891
Variables left out of the equation		
GRE – Composite Score	0.000	

As expected, multicollinearity was evident. There were three reasons for the excessively low tolerance and high VIF figures:

1. Biderman's Formula Score contains all three GRE (PR) component scores
2. GRE Composite Score includes the three GRE component scores
3. Total number of advanced courses includes athletic training and advanced math and science course, so only the number of advanced science courses was used

An examination of several combinations of variables led to a decision that the three predictors listed above be dropped from the final combination of discrete and continuous variables. The final set of predictor variables selected is shown in Table 4.5.

Table 4.5 Multicollinearity analysis results for discrete and continuous predictors retained

	Multicollinearity Statistics	
	Tolerance	**VIF**
GPA at the end of the first-yr	0.608	1.646
Number of Adv Math & Science Courses	0.844	1.185
GREv	0.589	1.698
GREq	0.495	2.021
GREwr	0.735	1.360

Next, the multi-level discrete and continuous variables were dichotomized through ROC analysis. The results of the multicollinearity assessment of the eight dichotomized variables, two of which were dichotomous at the outset, are presented in Table 4.6.

Table 4.6 Multicollinearity analysis results for dichotomous predictor variables

	Multicollinearity Statistics	
	Tolerance	**VIF**
Advanced Math & Science Courses ≥ 3	0.544	1.837
GPA end of first-year ≥ 3.45	0.704	1.420
GREv ≥ 145.5 (PR ≥ 26)	0.490	2.040
GREq ≥ 143.5 (PR ≥ 16.5)	0.634	1.578
GREwr ≥ 3.25 (PR ≥ 24.5)	0.815	1.227
Biderman's Formula Score ≥ 420.5	0.416	2.406
Physics: 1 = Yes; 0 = No	0.490	2.039
Calculus: 1 = Yes; 0 = No	0.689	1.450

The predictor variables outlined in Table 4.6 above were included in a logistic regression analysis to determine the best combination of variables to predict success on a student's first-attempt on the BOC exam.

Logistic Regression Analysis

All of the dichotomized predictor variables were entered into a logistic regression analysis with "first-attempt pass – Yes or No, on the BOC exam" as the outcome variable. Although multicollinearity testing did not reveal overlap between Biderman's Formula Score and the GRE, or between Advanced Courses and either Physics or Calculus, adjusted OR values were much smaller than the OR values derived from the separate univariable analysis. The results of the initial logistic regression analysis are displayed in Table 4.7.

Table 4.7 Logistic regression analysis results including all potential predictors of first-attempt BOC exam success

		Adjusted OR	95% C.I. Lower	95% C.I. Upper
Step 1	Advanced Math & Science Courses ≥ 3	2.361	0.494	11.293
	gGPA 1stYr. ≥ 3.45	7.564	1.845	31.007
	GREv ≥ 145.5 (PR ≥ 26)	3.385	0.677	16.915
	GREq ≥ 143.5 (PR ≥ 16.5)	5.016	1.115	22.563
	GREwr ≥ 3.25 (PR ≥ 24.5)	2.290	0.538	9.744
	Biderman Formula Score ≥ 420.5	0.555	0.095	3.234
	Physics Yes or No	0.836	0.155	4.506
	Calculus Yes or No	0.154	0.024	0.979
	Constant	0.220		
Step 2	Advanced Math & Science Courses ≥ 3	2.175	0.555	8.520
	gGPA 1stYr. ≥ 3.45	7.552	1.847	30.885
	GREv ≥ 145.5 (PR ≥ 26)	3.300	0.675	16.136
	GREq ≥ 143.5 (PR ≥ 16.5)	4.719	1.176	18.932
	GREwr ≥ 3.25 (PR ≥ 24.5)	2.271	0.534	9.658
	Biderman Formula Score ≥ 420.5	0.582	0.106	3.191
	Calculus Yes or No	0.148	0.024	0.905
	Constant	0.217		
Step 3	Advanced Math & Science Courses ≥ 3	2.271	0.586	8.806
	gGPA 1stYr. ≥ 3.45	6.816	1.748	26.572
	GREv ≥ 145.5 (PR ≥ 26)	2.435	0.699	8.481
	GREq ≥ 143.5 (PR ≥ 16.5)	4.246	1.132	15.928
	GREwr ≥ 3.25 (PR ≥ 24.5)	2.136	0.511	8.934
	Calculus Yes or No	0.152	0.025	0.923

	Constant	0.229		
Step 4	Advanced Math & Science Courses ≥ 3	1.991	0.528	7.502
	gGPA 1stYr. ≥ 3.45	7.148	1.860	27.464
	GREv ≥ 145.5 (PR ≥ 26)	2.917	0.884	9.628
	GREq ≥ 143.5 (PR ≥ 16.5)	4.560	1.245	16.696
	Calculus Yes or No	0.161	0.026	0.985
	Constant	0.363		
Step 5	gGPA 1stYr. ≥ 3.45	6.538	1.746	24.489
	GREv ≥ 145.5 (PR ≥ 26)	2.984	0.905	9.843
	GREq ≥ 143.5 (PR ≥ 16.5)	4.454	1.236	16.047
	Calculus Yes or No	0.244	0.049	1.199
	Constant	0.450		

This model produced five steps, which step five appeared to provide the best fit with a Nagelkerke R^2 at 0.386. However, at step five, the adjusted OR for Calculus was below 1.0. (Note: SPSS produced only five steps for this logistic regression analysis.) Due to potential conflict between Biderman's Formula Score and the GRE component scores and between the Advance Math & Science Courses and Physics and Calculus, Biderman's Formula Score and the individual courses were removed from the model and the logistic regression analysis was repeated. The results of this logistic regression analysis are shown in Table 4.8.

Table 4.8 Second logistic regression analysis results (Biderman's Formula Score. Physics and Calculus removed) for prediction of first-attempt BOC exam success

		Adjusted OR	95% C.I.	
			Lower	Upper
Step 1	Advanced Math & Science Courses ≥ 3	2.054	0.619	6.817
	gGPA 1stYr. ≥ 3.45	4.597	1.389	15.211
	GREv ≥ 145.5 (PR ≥ 26)	2.336	0.714	7.637
	GREq ≥ 143.5 (PR ≥ 16.5)	2.521	0.706	9.003
	GREwr ≥ 3.25 (PR ≥ 24.5)	1.945	0.484	7.814
	Constant	0.246		
Step 2	Advanced Math & Science Courses ≥ 3	1.930	0.590	6.318
	gGPA 1stYr. ≥ 3.45	4.834	1.487	15.718
	GREv ≥ 145.5 (PR ≥ 26)	2.695	0.862	8.428
	GREq ≥ 143.5 (PR ≥ 16.5)	2.775	0.792	9.719
	Constant	0.366		
Step 3	gGPA 1stYr. ≥ 3.45	4.432	1.404	13.988
	GREv ≥ 145.5 (PR ≥ 26)	2.668	0.853	8.343
	GREq ≥ 143.5 (PR ≥ 16.5)	3.494	1.066	11.448
	Constant	0.471		

Step three of the analysis appeared to provide the best fit with a Nagelkerke R^2 at 0.353. The set of three predictor variables included gGPA at the end of the first year ≥ 3.45, GREv ≥ 145.5 (PR ≥ 26), and GREq ≥ 143.5 (PR ≥ 16.5).

A second logistic regression analysis was performed that included the following dichotomized predictor variables: advanced math and science courses, gGPA at the end of the first-year, the Biderman's Formula Score, the student taking physics as an undergraduate, and the student taking calculus as an undergraduate, with “first-attempt pass – Yes or No, on the BOC exam” as the outcome variable (Table 4.9).

Table 4.9 Logistic regression analysis results (including Biderman’s Formula Score) for prediction of first-attempt BOC exam success

			95% C.I.	
		Adjusted OR	**Lower**	**Upper**
Step 1	Advanced Math & Science Courses ≥ 3	1.315	.313	5.524
	gGPA 1stYr. ≥ 3.45	8.692	2.345	32.220
	Biderman’s Formula Score ≥ 420.5	2.560	.787	8.332
	Physics Yes or No	2.310	.523	10.205
	Calculus Yes or No	.214	.039	1.165
	Constant	.614		
Step 2	gGPA 1stYr. ≥ 3.45	8.667	2.341	32.082
	Biderman’s Formula Score ≥ 420.5	2.565	.791	8.317
	Physics Yes or No	2.679	.748	9.604
	Calculus Yes or No	.225	.042	1.201
	Constant	.632		
Step 3	gGPA 1stYr. ≥ 3.45	7.812	2.152	28.356
	Biderman’s Formula Score ≥ 420.5	2.483	.780	7.907
	Calculus Yes or No	.422	.103	1.732
	Constant	.935		
Step 4	gGPA 1stYr. ≥ 3.45	5.783	1.866	17.923
	Biderman’s Formula Score ≥ 420.5	2.336	.746	7.320
	Constant	.901		
Step 5	gGPA 1stYr. ≥ 3.45	8.193	2.884	23.274
	Constant	1.133		

Step four of the analysis provided the best fit with a Nagelkerke R^2 at 0.263. The final prediction model had only two predictors: gGPA at the end of the first year ≥ 3.45 and a Biderman's Formula Score ≥ 420.5.

Prediction Model

The multiple analyses yielded two potential models for prediction of passing the BOC exam on the first-attempt: a three-factor model that included a gGPA at the end of the first-year ≥ 3.45, GREv ≥ 145.5 (PR ≥ 26), and GREq ≥ 143.5 (PR ≥ 16.5); and a two-factor model that included a gGPA at the end of the first year ≥ 3.45, and having a Biderman's Formula Score of ≥ 420.5.

For each prediction model, the sum of the number of positive predictor variables for each subject was calculated, and an ROC analysis was performed to identify the number of positive factors that provided the best balance of Sn and Sp for prediction of first-attempt BOC exam success. The results of this analysis are provided in Figure 4.2 and Table 4.10.

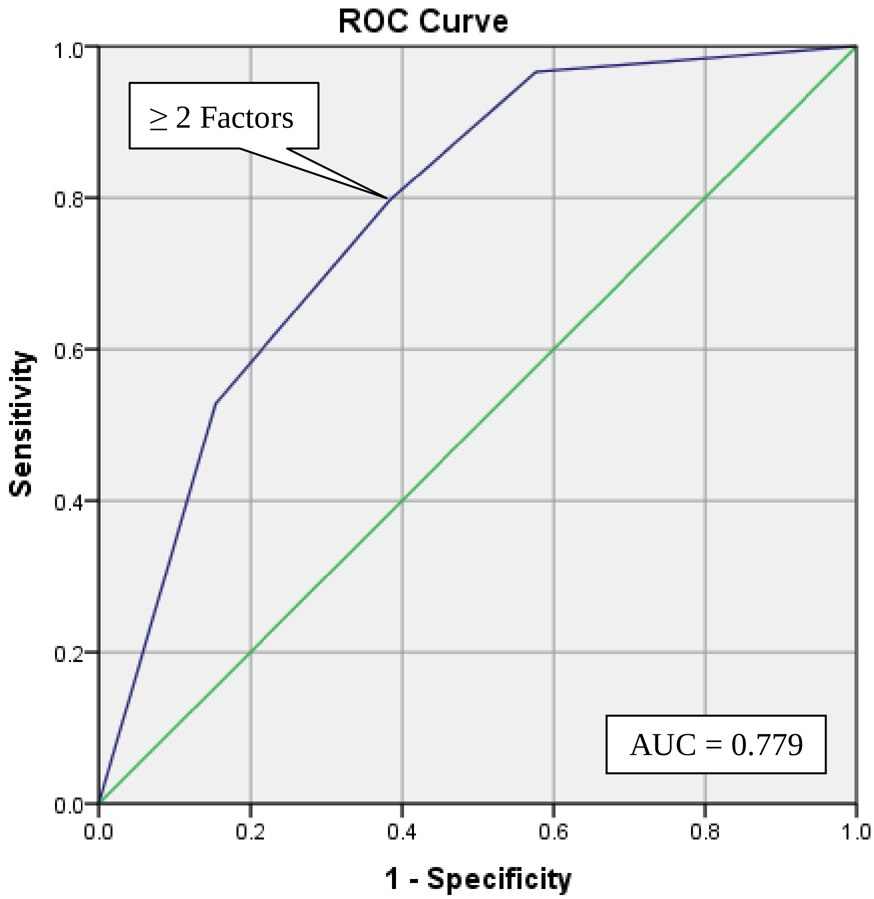

Figure 4.2 ROC curve with identification of the optimum cut-point for the number of positive factors (including GREv and GREq scores) for prediction of first-attempt BOC exam success

Table 4.10 Number of factors (including GREv and GREq scores) for prediction of first-attempt BOC exam success

	First-attempt Pass on the BOC exam	
	Yes	**No**
≥ 2 Factors	71	10
< 2 Factors	18	16
Fisher's Exact Test (one-sided) *p* < 0.001		
Sn = 0.80 (95% CI: 0.70 – 0.87)	Sp = 0.62 (95% CI: 0.43 – 0.78)	
Youden's Index = 0.413		
OR = 6.31 (95% CI: 2.46 – 16.23)	RFS = 1.66 (95% CI: 1.35 – 2.03)	

A GATP student who had ≥ 2 positive factors, (gGPA at the end of the first year ≥ 3.45, GREv ≥ 145.5 (PR ≥ 26), or GREq ≥ 143.5 [PR ≥ 16.5]), had 6.31 times greater odds of first-attempt BOC exam success than the odds for someone who had none or only one of the three factors. The relative frequency of success indicates the probability of a student passing the BOC exam on the first-attempt with any two or more of these factors is slightly more than one and half times the probability of a student who has less than two of these factors.

The percentages of successful GATP students according to the number of positive factors are presented in Table 4.11.

Table 4.11 Specific number of factors (including GREv and GREq scores) for prediction of first-attempt pass – Yes or No, on the BOC exam

	First-attempt Pass on the BOC exam				
Number of Factors	**Yes**	**No**	**Total**	**Percentage**	**Percentage above/ below cut point**
0	3	11	14	21.43%	18/34 = 52.94%
1	15	5	20	75.00%	
2	24	6	30	80.00%	71/81 = 87.65%
3	47	4	51	92.16%	
Total	89	26	115	77.39%	

Among students who had two or more positive factors, 87.65% passed the BOC exam on the first-attempt. Of the students who had one or

none of the positive factors only 52.94% achieved BOC exam eligibility and passed on the exam on the first-attempt.

A two-factor model including Biderman's Formula Score is presented in Figure 4.3 and Table 4.12.

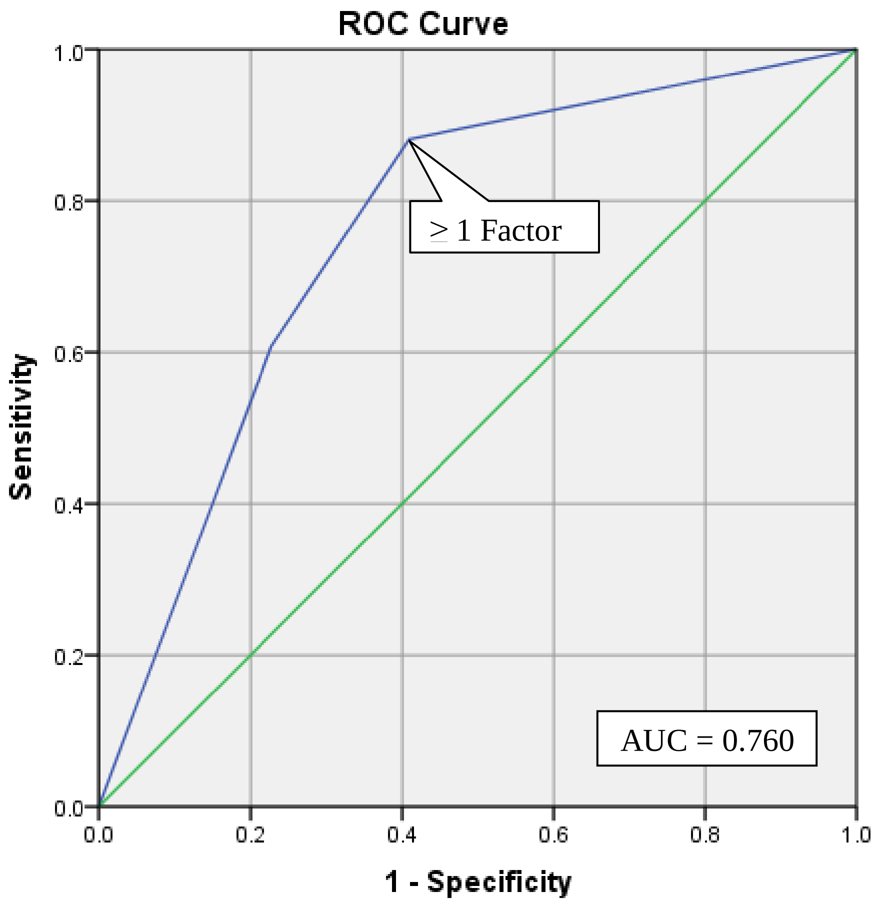

Figure 4.3 ROC curve with identification of the optimum cut-point for the number of positive factors (including Biderman's Formula Score) as a predictor of first-attempt BOC exam success

Table 4.12 Number of factors (including Biderman's Formula Score) for prediction of first-attempt pass – Yes or No, on the BOC exam

	First-attempt Pass on the BOC exam	
	Yes	No
≥ 1 Factor	74	9
No Factors	10	13
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.88 (95% CI: 0.80 – 0.93)	Sp = 0.59 (95% CI: 0.39 – 0.77)	
Youden's Index = 0.498		
OR = 10.69 (95% CI: 3.64 – 31.16)	RFS = 2.05 (95% CI: 1.67 – 2.51)	

For the two-factor model, a GATP student who had at least one positive factor, (either gGPA at the end of the first year of ≥ 3.45, or Biderman's Formula Score of ≥ 420.5) had 10.69 times greater odds of BOC exam success on the first-attempt than the odds for someone who had neither of the two factors. The relative frequency of GATP success indicates the probability of a student being successful in the GATP with one or more factors is slightly greater than twice the probability of a student with none of the positive factors. The percentages of successful GATP students according to the number of positive factors are presented in Table 4.13.

Table 4.13 Specific number of factors (including Biderman's Formula Score) for prediction of first-attempt pass – Yes or No, on the BOC exam

Number of Factors	First-attempt Pass on the BOC exam: Yes	No	Total	Percentage	Percentage above/ below cut point
0	10	13	23	43.48%	
1	23	4	27	85.19%	74/83 = 89.16%
2	51	5	56	92.73%	
Total	89	26	115	77.39%	

Among students who had one or more positive factors, 89.16% passed the BOC exam on the first-attempt. Of the students who had none of the positive factors only 43.48% achieved BOC exam eligibility and passed on the exam on the first-attempt (Table 4.13).

Prediction of Success in GATP

The second purpose of this study was to utilize the results of the first analysis to identify program applicant characteristics that are most likely to predict both academic success in the graduate professional program and subsequent first-attempt success on the BOC exam. Because first-year gGPA (≥ 3.45) was found to be the strongest predictor of first-attempt BOC exam success, it was selected as the outcome variable for GATP success. When selecting the most qualified candidates for a GATP, the goal is to recruit students who most likely to pass the BOC exam on the first-attempt.

To create a prediction model, the initial step is to identify all possible predictor variables that might have an association with the outcome variable (Bruce & Wilkerson, 2010a; Childs &

Cleland, 2006). A list of 39 potential predictor variables is presented in Table 4.14. The following were multi-level discrete variables: number of advanced courses, number of athletic training courses, and number of advanced science courses. The following were continuous variables: Institution ACT mean/median or SAT mean/median, uGPA, GRE component score, GREv, GREq, GREwr, and Biderman's Formula Score.

Table 4.14 Potential predictor variables analyzed

- Academic Profile of Undergraduate Institution (APUI)
  - Undergraduate institution SAT mean/median
  - Undergraduate institution ACT mean/median
  - Undergraduate institution SAT 75th percentile
  - Undergraduate institution ACT 75th percentile
  - Undergraduate institution SAT 80th percentile
  - Undergraduate institution ACT 80th percentile

- Basic Carnegie classification categories
  - Bachelors Only
  - Bach & Masters
  - Doctorate/Research
  - Research Intensive

- Undergraduate institution size and setting:
  - Large (10,000+ undergraduates)
  - Medium (3,000-9,999 undergraduates)
  - Small (<1,000-2,999 undergraduates)

- Advanced math and science courses
  - Number of advanced science courses
  - Any advanced biology
  - Any advanced chemistry
  - Biomechanics
  - Calculus
  - Pathophysiology
  - Physics

- Athletic training courses
  - Number of athletic training courses
  - Basic athletic training or Care & Prevention courses
  - Advanced athletic training courses

- Advanced math, science, and athletic training courses
  - Total number of advanced courses

- uGPA

- GRE Scores
  - GRE Composite
  - GREq
  - GREv
  - GREwr

- Biderman's Formula Score

For each of the multi-level discrete and continuous variables, an ROC analysis was performed to determine the optimum cut-point for dichotomization. Each of the dichotomized predictor variables was analyzed by univariable 2 X 2 cross-tabulation, which included the calculation of Sn, Sp, OR, RFS and the *p*-value for Fisher's Exact Test (one-sided).

Determination of Academic Profile of Undergraduate Institution

The variable Academic Profile of Undergraduate Institution (APUI) was quantified by examining each institution's reported ACT mean or median value and/or SAT mean or median value, and the 75th and 80th percentiles for these variables. The descriptive statistics related to the APUI are presented in Table 4.15.

Table 4.15 Descriptive statistics and summary of univariable analysis results for undergraduate institutions (N = 194) as potential predictors of first-year gGPA ≥ 3.45 relating to APUI for students admitted to GATP

Academic Profile of Undergraduate Institution	**[a]Mean (± sd)**	**Cut-point**	**Sn**	**1 - Sp**	**Sp**	**Youden's Index**	**AUC**	**OR**	**RFS**	**Fisher's Exact Test (one-sided) *p*-value**
Institution ACT mean/median (N = 110)	1128.3 (± 116.88)	≥ 25.5	0.48	0.14	0.86	0.341	0.710	5.82	1.54	0.001
Institution SAT mean/median (N = 121)	24.45 (± 2.82)	≥ 1132.5	0.61	0.29	0.71	0.318	0.697	3.78	1.44	0.003

[a]This is the mean (± sd) for all of the undergraduate institutions represented of students admitted to the GATP

A series of ROC analyses and corresponding 2 X 2 cross-tabulation tables were produced (The individual ROC analyses and 2 X 2 cross-tabulation tables for individual predictors are provided in Appendix B).

The summary of APUI statistics for potential predictor variables is presented in Table B.2, which are listed in order of the odds ratio magnitude. A list of the undergraduate colleges and universities with their respective ACT and SAT mean/median scores is provided in Appendix C.

To determine the best combination of reported Institution ACT and SAT scores to define high versus low APUI, various pairings of values were assessed through 2 X 2 cross-tabulation tables. The analysis results for the eight pairings are presented in Appendix D. To be classified as high APUI a school had a reported ACT mean/median of ≥ 25.5 or SAT mean/median of ≥ 1132.5. A college or university that reported their scores below the identified values were determined as low APUI. Since all of the pairings had relatively similar Sn, Sp, OR, RFS, and Fisher's Exact Test (one-sided) it was difficult to select which combination of ACT and SAT scores should be used for further consideration in the prediction model. The result of the analysis of either Institution ACT mean/median ≥ 25.5 or Institution SAT mean/median ≥ 1132.5 is shown Table 4.16. Ultimately the pairing selected provided the best balance between Sn and Sp, and the absolute mean/median figures were easier to locate on a college/university web site compared to the percentile ranks of the undergraduate athletic training students' institutions.

Table 4.16 Institution SAT mean/median for prediction of first-year gGPA ≥ 3.45

	First-year gGPA ≥ 3.45	**First-year gGPA < 3.45**
Either Institution ACT mean/median ≥ 25.5 or Institution SAT mean/median ≥ 1132.5	52	8
Neither Institution ACT mean/median ≥ 25.5 nor Institution SAT mean/median ≥ 1132.5	41	34
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.56 (95% CI: 0.46 – 0.66)	Sp = 0.81 (95% CI: 0.67 – 0.90)	
OR = 5.39 (95% CI: 2.25 – 12.89)	RFS = 1.59 (95% CI: 1.29 – 1.94)	

The OR of 5.39 for APUI classification met the criterion for inclusion in a multivariable analysis of potential predictors. The results of the univariable analyses for the potential predictors of first-year success gGPA (≥ 3.45) are presented in Appendix E, and summarized in Table 4.17, (variables are listed in order of the OR magnitude). The policy of the UTC Graduate School is to determine uGPA by combining all courses taken at all undergraduate institutions, which is the method utilized to determine each student's uGPA for this study.

Table 4.17 Summary of univariable analysis results for prediction of first-year gGPA ≥ 3.45

Variable - 3.45 gGPA	Cut-point	Sn	1 - Sp	Sp	Youden's Index	AUC	OR	RFS	Fisher's Exact Test (one-sided)
Biderman's Formula Score	458.45	0.61	0.09	0.91	0.528	0.816	16.94	1.84	0.001
GREq	141.5	0.90	0.47	0.53	0.430	0.772	10.49	2.66	0.001
*Calculus Yes or No	–	0.44	0.07	0.93	–	–	10.06	1.62	0.001
GRE - Composite	292.5	0.70	0.24	0.76	0.465	0.795	7.60	1.79	0.001
GREv	150.5	0.47	0.11	0.90	0.363	0.754	7.48	1.54	0.001
uGPA	3.18	0.71	0.33	0.67	0.380	0.715	4.71	1.67	0.001
Number of adv math & science courses	4	0.36	0.14	0.86	0.212	0.632	3.30	1.35	0.009
Number of adv courses	5	0.38	0.19	0.81	0.186	0.624	2.56	1.29	0.045
GREwr	3.75	0.66	0.46	0.54	0.202	0.648	2.30	1.28	0.044
*Graduated from a Research Intensive Institution Yes or No	–	0.46		0.67	–	–	1.69	1.17	0.121
*Physics Yes or No	–	0.58		0.52	–	–	1.52	1.14	0.173

Note. For further consideration a variable had to have an OR of ≥ 1.50 (Hosmer & Lemeshow, 2000) and a Fisher's Exact Test (one-sided) *p*-value of ≤ 0.20 (Bruce & Wilkerson, 2010a; Kuijpers et al., 2006; Teyhen et al., 2007)

*Dichotomous variables

A summary of the univariable analysis results for potential predictors that did not meet the criterion for inclusion in the multivariable analysis is presented in Table 4.18.

Table 4.18 Predictor variables eliminated from further consideration

	OR	**95% Confidence Interval**	**Fisher's Exact Test *p*-value (1-sided)**
<u>Carnegie classifications</u>			
Bachelors only	1.520	CI: 0.56 to 4.13	0.284
Bachelors and Master	0.773	CI: 0.37 to 1.62	0.310
Doctoral research	1.294	CI: 0.62 to 2.71	0.310
Graduate Program	1.322	CI: 0.58 to 3.00	0.322
Number of Athletic Training Courses (≥ 4 courses)	2.11	CI. 0.57 to 7.85	0.200
Public-Private	0.605	CI: 0.26 to 1.39	0.161
Residency (In-state vs. Out-of-state)	0.541	CI: 0.24 to 1.22	0.100
Size & Setting - small	1.540	CI: 0.63 to 3.77	0.234
Size & Setting - medium	0.474	CI: 0.21 to 1.08	0.590
Size & Setting - large	1.305	CI: 0.63 to 2.70	0.298
Took Basic AT courses	0.710	CI: 0.34 to 1.48	0.234
Took Advanced AT courses	1.055	CI: 0.44 to 2.55	0.548
Took biomechanics	1.418	CI: 0.66 to 3.04	0.240
Took advanced chemistry	1.403	CI: 0.66 to 2.96	0.242
Took advanced biology	1.276	CI: 0.60 to 2.71	0.329

Although ≥ 4 athletic training courses demonstrated an OR > 2.0, the lower limit of its 95% confidence interval was < 1.0 (0.57). Thus this potential predictor was dropped from further consideration.

Multicollinearity

A series of linear regression analyses were performed to assess multicollinearity among continuous variables, which included: the number of advanced math and science courses, total number of advanced courses, APUI (Institution ACT mean/median or Institution SAT mean/median), uGPA, GRE Composite score, GREv, GREq, GREwr, and Biderman's Formula Score. The analysis results from are presented in Table 4.19.

Table 4.19 Results for assessment of multicollinearity among potential predictors of first-year gGPA $\geq$ 3.45

	Multicollinearity Statistics	
	Tolerance	**VIF**
APUI	0.680	1.471
Number of adv math & science courses	0.174	5.757
Total number of adv courses (AT + Adv Science)	0.182	5.504
GREv	0.083	11.998
GREq	0.095	10.475
GREwr	0.066	15.089
uGPA	0.043	23.331
Biderman's Formula Score	0.009	107.068
<u>Variables left out of the equation</u>		
GRE – Composite score	0.000	

There were three reasons for the excessively low tolerance and high VIF values that were obtained:

1. Biderman's Formula Score contains all three GRE PR scores
2. The GRE Composite score includes the three parts of the GRE
3. Total number of advanced courses includes the number of advanced science courses

Through trial and error various combinations of multi-level discrete and continuous variables were selected (Table 4.20 and Table 4.21).

Table 4.20 Multicollinearity analysis results for seven-variable set of potential predictors (including GRE scores) of first-year gGPA ≥ 3.45

	Multicollinearity Statistics	
	Tolerance	**VIF**
APUI	0.675	1.481
Number of adv science courses	0.174	5.750
Total number of adv courses (AT + Adv Science)	0.180	5.551
GREv	0.517	1.934
GREq	0.498	2.007
GREwr	0.769	1.300
uGPA	0.836	1.196

Tolerance and VIF scores improved for uGPA and GRE scores, when the set of variables was reduced from 8 (Table 4.21) to seven (Table 4.22) by removal of Biderman's Formula Score, but there was still overlap between the variables. The analysis was repeated after removal of the "Total number of adv courses" variable (Table 4.23).

Table 4.21 Multicollinearity analysis results for six-factors set of potential predictors (including GRE scores) of gGPA ≥ 3.45

	Multicollinearity Statistics	
	Tolerance	**VIF**
APUI (Institution ACT + Institution SAT mean/median)	0.711	1.407
Number of adv science courses	0.830	1.204
GREv	0.522	1.917
GREq	0.505	1.982
GREwr	0.782	1.279
uGPA	0.844	1.185

This six-factor model demonstrates acceptable tolerance and VIF values. Multicollinearity analysis was repeated performed for a set of eight dichotomized variables (Table 4.22).

Table 4.22 Multicollinearity analysis results for an eight-factor set of dichotomized potential predictors (including GRE scores) of first-year gGPA ≥ 3.45

	Multicollinearity Statistics	
	Tolerance	**VIF**
High APUI	0.587	1.703
Number of Advanced Math & Science Courses ≥ 4	0.767	1.304
uGPA ≥ 3.18	0.878	1.139
Physics: 1 = Yes; 0 = No	0.672	1.487
Calculus: 1 = Yes; 0 = No	0.575	1.739
Research Intensive = 1; Others = 0	0.783	1.277
GREv ≥ 150.5 (PR ≥ 46.5)	0.759	1.317
GREq ≥ 141.5 (PR ≥ 12)	0.768	1.303
GREwr ≥ 3.75 (PR ≥ 44.5)	0.862	1.160

This eight-factor model demonstrates acceptable tolerance and VIF values. Results for assessment of multicollinearity among potential predictors of first-year gGPA ≥ 3.45 found excessively low tolerance and high VIF values (Table 4.19). Because Biderman's Formula Score contains all three GRE PR scores, it was dropped from this specific analysis.

A second series of analyses were performed to assess multicollinearity among continuous and multi-discrete variables, which included: the number of advanced math and science courses, APUI (Institution ACT mean/median or Institution SAT mean/median), uGPA, and Biderman's Formula Score. The analysis results from are presented in Table 4.23.

Table 4.23 Results for assessment of multicollinearity among potential predictors (including Biderman's Formula Score) of first-year gGPA $\geq$ 3.45

	Multicollinearity Statistics	
	Tolerance	**VIF**
APUI	0.738	1.353
Number of adv math & science courses	0.892	1.122
uGPA	0.463	2.159
Biderman's Formula Score	0.388	2.577

This four-factor model demonstrates acceptable tolerance and VIF values. Multicollinearity analysis was then performed for a set of seven dichotomized variables (Table 4.24).

Table 4.24 Multicollinearity analysis results for seven-factor set of dichotomized variables (including Biderman's Formula Score) for prediction of first-year gGPA ≥ 3.45

	Multicollinearity Statistics	
	Tolerance	**VIF**
High APUI	0.635	1.575
Total Advanced Courses ≥ 5	0.801	1.249
uGPA ≥ 3.18	0.804	1.243
Biderman's Formula Score ≥ 458.45	0.686	1.457
Physics: 1 = Yes; 0 = No	0.735	1.360
Calculus: 1 = Yes; 0 = No	0.574	1.743
Research Intensive = 1; Others = 0	0.831	1.204

Logistic Regression Analysis

GRE Model

Because two possible prediction models were created to forecast gGPA at the end of the first year, two separate logistic regression analyses were performed. The first analysis included the GRE component scores with five other dichotomized predictor variables. The variables included in this analysis were: High APUI, uGPA ≥ 3.18, ≥ 4 advance math & science courses, GREv ≥ 150.5 (PR ≥ 46.5), GREq ≥ 141.5 (PR ≥ 12.0), GREwr ≥ 3.75 (PR ≥ 44.5), graduated from a research intensive institution, took physics as an undergraduate, and took calculus as an undergraduate (Table 4.25).

Table 4.25 Logistic regression analyses of nine variables for prediction of first-year gGPA ≥ 3.45

		Adjusted OR	95% C.I. Lower	95% C.I. Upper
Step 1	High APUI	0.703	0.182	2.708
	Number of math & science courses ≥ 4	1.870	0.314	11.136
	uGPA ≥ 3.18	7.661	2.303	25.485
	GREv ≥ 150.5 (PR ≥ 46.5)	3.137	0.730	13.489
	GREq ≥ 141.5 (PR ≥ 12)	7.041	1.848	26.827
	GREwr ≥ 3.75 (PR ≥ 44.5)	1.100	0.370	3.264
	Research Intensive = 1; Others = 0	2.054	0.593	7.121
	Physics: 1 = Yes; 0 = No	0.665	0.184	2.407
	Calculus: 1 = Yes; 0 = No	13.353	2.060	86.548
	Constant	0.081		
Step 2	High APUI	0.701	0.182	2.700
	Number of math & science courses ≥ 4	1.858	0.315	10.968
	uGPA ≥ 3.18	7.771	2.355	25.638
	GREv ≥ 150.5 (PR ≥ 46.5)	3.194	0.756	13.497
	GREq ≥ 141.5 (PR ≥ 12)	7.053	1.853	26.851
	Research Intensive = 1; Others = 0	2.101	0.622	7.097
	Physics: 1 = Yes; 0 = No	0.668	0.185	2.411
	Calculus: 1 = Yes; 0 = No	13.444	2.076	87.066
	Constant	0.084		
Step 3	Number of math & science courses ≥ 4	1.908	0.319	11.402
	uGPA ≥ 3.18	7.339	2.276	23.664
	GREv ≥ 150.5 (PR ≥ 46.5)	2.972	0.720	12.275
	GREq ≥ 141.5 (PR ≥ 12)	6.420	1.791	23.018
	Research Intensive = 1; Others = 0	1.942	0.599	6.296
	Physics: 1 = Yes; 0 = No	0.690	0.193	2.462
	Calculus: 1 = Yes; 0 = No	10.981	2.012	59.929
	Constant	0.086		

Step 4	Number of math & science courses ≥ 4	2.146	0.381	12.091
	uGPA ≥ 3.18	8.162	2.645	25.186
	GREv ≥ 150.5 (PR ≥ 46.5)	2.899	0.710	11.833
	GREq ≥ 141.5 (PR ≥ 12)	5.623	1.722	18.360
	Research Intensive = 1; Others = 0	1.904	0.587	6.176
	Calculus: 1 = Yes; 0 = No	9.336	1.917	45.472
	Constant	0.076		
Step 5	uGPA ≥ 3.18	7.300	2.477	21.510
	GREv ≥ 150.5 (PR ≥ 46.5)	2.650	0.665	10.564
	GREq ≥ 141.5 (PR ≥ 12)	6.442	2.052	20.225
	Research Intensive = 1; Others = 0	1.795	0.558	5.771
	Calculus: 1 = Yes; 0 = No	8.716	1.829	41.538
	Constant	0.085		
Step 6	uGPA ≥ 3.18	7.018	2.418	20.375
	GREv ≥ 150.5 (PR ≥ 46.5)	2.828	0.696	11.486
	GREq ≥ 141.5 (PR ≥ 12)	6.087	1.959	18.916
	Calculus: 1 = Yes; 0 = No	9.481	2.062	43.594
	Constant	0.104		
Step 7	uGPA ≥ 3.18	7.624	2.627	22.127
	GREq ≥ 141.5 (PR ≥ 12)	7.677	2.481	23.759
	Calculus: 1 = Yes; 0 = No	11.767	2.657	52.106
	Constant	0.101		

Step 7 produced the best model of fit, with a Nagelkerke R^2 of 0.493. The lower limit 95% CI for the adjusted OR was > 1.0 for all three variables: uGPA ≥ 3.18, GREq ≥ 141.5 (PR ≥ 12), and having taken calculus as an undergraduate.

A second logistic regression analysis was performed that included all of the dichotomized predictor variables, including Biderman's Formula Score, with gGPA at the end of the first year ≥ 3.45 as the outcome variable. The predictor variables included the following: High APUI, uGPA $\geq$ 3.18, $\geq$ 4 advance math & science courses, Biderman's Formula Score $\geq$ 458.45, graduated from a research intensive institution, took physics as an undergraduate student, and took calculus as an undergraduate. The analysis generated a model of five steps from the logistic regression analysis. All of the steps and the adjusted OR and the associated 95% confidence interval are shown in Table 4.26. Step 5 produced the best model of fit and had a Nagelkerke R^2 of 0.436.

Table 4.26 Logistic regression analysis results including Biderman's Formula Score as predictors of first-year gGPA $\geq$ 3.45

		Adjusted OR	95% C.I.	
			Lower	Upper
Step 1	High APUI	1.490	0.465	4.777
	Number of math & science courses $\geq$ 4	2.250	0.619	8.176
	uGPA $\geq$ 3.18	3.211	1.218	8.466
	Biderman's Formula Score $\geq$ 458.45	7.631	1.959	29.733
	Research Intensive = 1; Others = 0	0.901	0.334	2.434
	Physics: 1 = Yes; 0 = No	0.580	0.208	1.622
	Calculus: 1 = Yes; 0 = No	6.228	1.228	31.580
	Constant	0.456		
Step 2	High APUI	1.469	0.463	4.657
	Number of math & science courses $\geq$ 4	2.240	0.617	8.130
	uGPA $\geq$ 3.18	3.248	1.239	8.518
	Biderman's Formula Score $\geq$ 458.45	7.572	1.949	29.420
	Physics: 1 = Yes; 0 = No	0.582	0.208	1.627
	Calculus: 1 = Yes; 0 = No	6.175	1.218	31.296
	Constant	0.440		
Step 3	Number of math & science courses $\geq$ 4	2.269	0.627	8.208
	uGPA $\geq$ 3.18	3.348	1.280	8.752
	Biderman's Formula Score $\geq$ 458.45	8.165	2.132	31.261
	Physics: 1 = Yes; 0 = No	0.561	0.203	1.555
	Calculus: 1 = Yes; 0 = No	7.888	1.844	33.732
	Constant	0.469		
Step 4	Number of math & science courses $\geq$ 4	1.890	0.550	6.496
	uGPA $\geq$ 3.18	3.487	1.341	9.066
	Biderman's Formula Score $\geq$ 458.45	7.745	2.052	29.235
	Calculus: 1 = Yes; 0 = No	6.177	1.551	24.598
	Constant	0.382		
Step 5	uGPA $\geq$ 3.18	3.180	1.249	8.093
	Biderman's Formula Score $\geq$ 458.45	8.331	2.221	31.249
	Calculus: 1 = Yes; 0 = No	7.113	1.822	27.770
	Constant	0.437		

The final three predictor variables were uGPA ≥ 3.18, Biderman's Formula Score ≥ 458.45, and took calculus as an undergraduate.

Interaction Effects

Because two models were able to predict gGPA at the end of the first year in the GATP, separate analyses were conducted to assess any interaction effects. The first logistic regression model included the GREq scores. The univariable odds ratio and multivariable adjusted odds ratio for each of the predictor variables is shown in Table 4.27.

Table 4.27 Comparison of odds ratios for predictor variables

	Univariable OR	**Multivariable Adj OR**
uGPA	4.71 (95% CI: 2.17 – 10.23)	7.62 (95% CI: 2.63 – 22.13)
GREq	10.49 (95% CI: 4.11 – 26.78)	7.68 (95% CI: 2.48 – 23.76)
Calculus	10.06 (95% CI: 2.90 – 34.86)	11.77 (95% CI: 2.66 – 52.11)

The existence of an interaction between uGPA and GREq is suggested by the differences between the univariable odds ratio and the corresponding multivariable adjusted odds ratio, whereas there was relatively little change between the two odds ratios for taking calculus.

The interaction pairings studied were: GREq X uGPA; uGPA X Calculus; GREq X Calculus. Each interaction pairing was examined for prediction of success, (success = gGPA ≥ 3.45 at the end of the first year). Each possible interaction was examined three ways:

1. 2 X 2 cross-tabulation tables to calculate the Sn, Sp, OR, RFS and Fisher's Exact Test (one-sided).
2. Stratified analysis and graphic representation of the interaction
3. Stratum-specific odd ratios were compared to the Mantel-Haenszel OR estimate and the Breslow-Day test was done to confirm homogeneity of the stratum-specific ORs.

The next series of tables and figures demonstrate the nature of the interactive relationship between GREq and uGPA (Tables 4.28 to 4.30 and Figure 4.4).

Table 4.28 A student with a combination of a high uGPA ($\geq$ 3.18) and a high GREq ($\geq$ 141.5 [PR $\geq$ 12]) for prediction of first-year gGPA $\geq$ 3.45

	First-year gGPA of ≥ 3.45	**First-year gGPA of < 3.45**
Both factors, uGPA X GREq	65	5
≤ 1 factor, either uGPA X GREq	29	35
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.69 (95% CI: 0.59 – 0.78)	Sp = 0.88 (95% CI: 0.74 – 0.95)	
OR = 15.69 (95% CI: 5.58 – 44.13)	RFS = 2.05 (95% CI: 1.67 – 2.51)	

A student who had both a high uGPA ($\geq$ 3.18) and a high GREq ($\geq$ 141.5 [PR $\geq$ 12.0]) had 15.69 times greater odds for success in the GATP than the odds for someone who had either one or none of the factors. The relative frequency of GATP success indicates the probability of a student being successful in the GATP who has both a high uGPA and a high GREq is slightly more than twice that for students who have only one or none of these factors.

Table 4.29 Stratified analysis of uGPA levels for association of GREq as a predictor of gGPA

uGPA ≥ 3.18				
	Success	**Not successful**	**Total**	**Percentage**
High GREq	65	5	70	93%
Low GREq	3	8	11	27%
OR = 34.67				

uGPA < 3.18				
	Success	**Not successful**	**Total**	**Percentage**
High GREq	20	13	33	61%
Low GREq	6	12	18	33%
OR = 3.07				

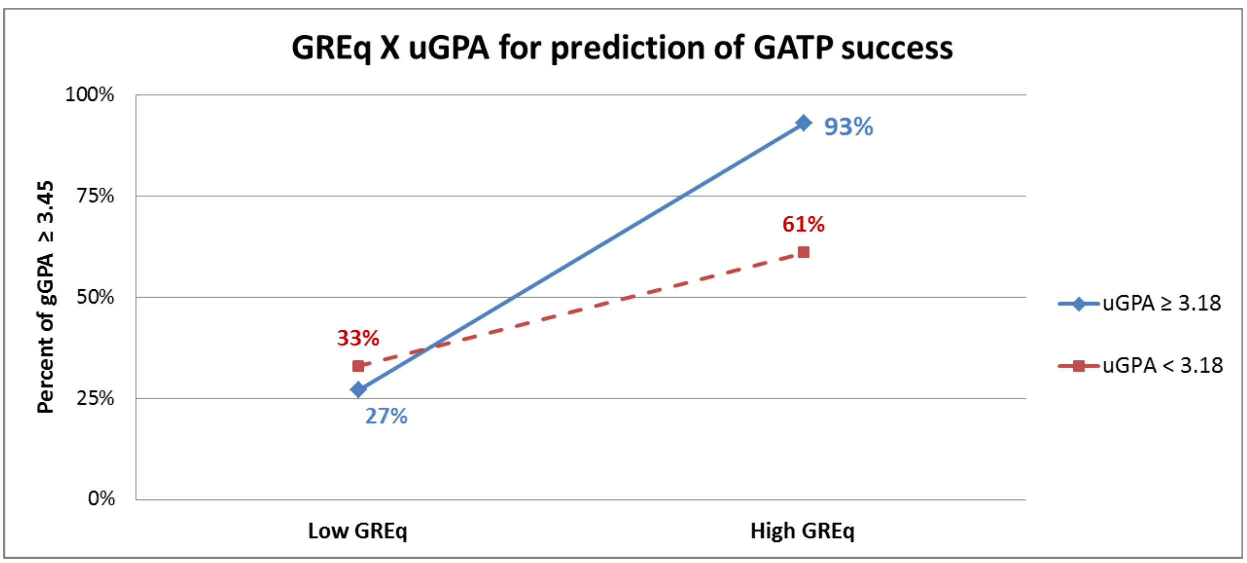

Figure 4.4 GREq X uGPA for prediction of GATP success

This possible interaction represents students with both a high GREq (≥ 141.5 [PR ≥ 12.0]) and a high uGPA (≥ 3.18) were 93% successful. Students with both a low GREq and a low uGPA had a low success rate (27%). A student who had a high uGPA and a high GREq had 34.67 times greater odds to be successful in the GATP than the odds for someone who had a high GPA and a low GREq. Conversely, a student who had a low uGPA and a high GREq had 3.07 times greater odds for GATP success than one who had a low uGPA and a low GREq. The OR indicates that a student who had a high GREq (≥ 141.5 [PR ≥ 12.0]) and had a high uGPA (≥ 3.18) had 34.67 times greater odds to be successful in the GATP than the odds for someone who had a low GREq and had a high uGPA. A student who had a high GREq and a low uGPA had 3.07 time greater odds to be successful in the GATP than the odds from someone who had a low GREq and a low uGPA.

Controlling for uGPA, the relationship between GREq and GATP success (gGPA at the end of the first year ≥ 3.45) was examined (Mantel-Haenszel OR_{est} = 6.49 [95% CI: 2.59 – 16.52]). There is a statistically significant association between GREq and GATP success (gGPA at the end of the first year ≥ 3.45) and high and low uGPA strata (≥ 3.18 OR = 34.67 [95% CI: 6.94 – 173.21]; < 3.18 OR = 3.08 [95% CI: 0.92 – 10.25]); Mantel-Haenszel $\chi^2(1) = 18.615$; ($p < 0.001$). The null hypothesis for the Breslow-Day test assumes that the odds ratios for GREq by GATP success (gGPA at the end of the first year ≥ 3.45) at the end of the first year is equivalent for uGPA categories. The Breslow-Day test for homogeneity found the odds ratios to be significantly different for the two strata, Breslow-Day $\chi^2(1) = 6.045$; ($p = 0.014$).

An examination of the univariable odds ratio and the multivariable adjusted odds ratio is shown in Table 4.30 for uGPA X GREq and Calculus.

Table 4.30 Univariable and multivariable comparison of odds ratio for the interaction of uGPA and GREq with taking calculus

	Univariable OR	**Multivariable Adj OR**
uGPA X GREq	15.69 (95% CI: 5.58 – 44.13)	16.80 (95% CI: 5.62 – 50.21)
Calculus	10.06 (95% CI: 2.90 – 34.86)	10.92 (95% CI: 2.85 – 41.89)

This table demonstrates that calculus appears to have an independent effect (10.06 – 10.92), but uGPA and GREq interact. A 2 X 2 analysis that does not include calculus (uGPA X GREq) generates an OR that is not very different from the multivariable adjusted OR derived from a logistic regression analysis that did include calculus (15.69 – 16.80).

The next series of tables and figures demonstrate the nature of the interactive relationship between uGPA and taking calculus for prediction of gGPA (Tables 4.31 to 4.33 and Figure 4.5).

Table 4.31 A student with a combination of a high uGPA (≥ 3.18) and took calculus as an undergraduate for prediction of first-year gGPA ≥ 3.45

	First-year gGPA of ≥ 3.45	**First-year gGPA of < 3.45**
Both factors, uGPA X Calculus (1)	27	1
≤ 1 factor, either uGPA X Calculus (0)	67	41
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.29 (95% CI: 0.206 – 0.386)	Sp = 0.98 (95% CI: 0.877 – 0.996)	
OR = 16.52 (95% CI: 2.163 – 1.905)	RFS = 1.55 (95% CI: 1.268 – 1.905)	

A student who had both a high uGPA (≥ 3.18) and had taken calculus as an undergraduate had 16.52 times greater odds to be successful in the GATP than the odds for someone who had only one or none of the factor. The relative frequency of GATP success indicates the probability of a student being successful in the GATP who had both a high uGPA and had taken calculus as an undergraduate is slightly more than one and half that for a students who only one or none of these factors. Please note the cell count of “1” is cause to interpret these results with skepticism since it weakens the overall analysis and results in highly unstable odd ratios (Hosmer & Lemeshow, 2000).

Table 4.32 Stratified analysis of uGPA for levels of association of calculus history as a predictor of gGPA

uGPA ≥ 3.18				
	Success	**Not successful**	**Total**	**Percentage**
Calculus - Yes	27	1	28	96%
Calculus - No	41	14	55	75%
OR = 9.22				

uGPA < 3.18				
	Success	**Not successful**	**Total**	**Percentage**
Calculus - Yes	14	2	16	88%
Calculus - No	12	25	37	32%
OR = 14.58				

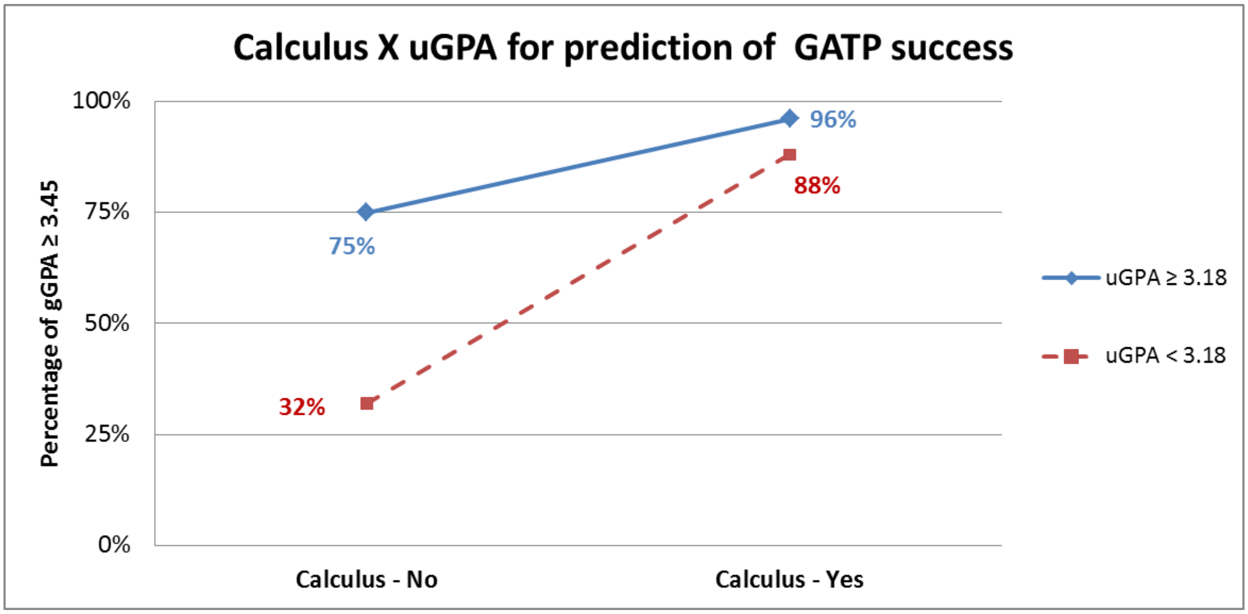

Figure 4.5 Calculus X uGPA for prediction of GATP success

The interaction indicates that students who took calculus had a high rate of success regardless of uGPA (uGPA ≥ 3.18 = 96%; uGPA < 3.18 = 88%). A student who took calculus and who had a high uGPA (≥ 3.18) had 9.22 times greater odds for success in the GATP than the odds for someone with a high uGPA, who did not take calculus. Students who took calculus, but had a low uGPA (< 3.18) had 14.58 times greater odds for success in the GATP than the odds for someone who had a low uGPA and did not take calculus.

Controlling for uGPA, the relationship between taking calculus and GATP success (gGPA at the end of the first year ≥ 3.45) was examined using a Mantel-Haenszel analysis (Mantel-Haenszel OR_{est} = 11.79 [95% CI: 3.71 – 44.12]). There is a statistically significant association between taking calculus and GATP success (gGPA at the end of the first year ≥ 3.45) and high and low uGPA strata (≥ 3.18 OR = 9.22 [95% CI: 1.15 – 74.25]; < 3.18 OR = 14.58 [95% CI: 2.85 – 74.71]); Mantel-Haenszel $\chi^2(1) = 16.76$; ($p < 0.001$). The null hypothesis for the Breslow-Day test assumes that the odds ratios for taking calculus by gGPA at the end of the first year is equivalent for uGPA categories. The Breslow-Day test for homogeneity found the odds ratios to be not significantly different from one another, Breslow-Day $\chi^2(1) = 0.119$; ($p = 0.730$). Please note the large confidence intervals are due to the low cell counts. Hosmer & Lemeshow (2000) suggest a minimum of five for each cell to have more reliable, valid, and stable model.

An examination of the univariable odds ratio and the multivariable adjusted odds ratio is shown in Table 4.33 for uGPA X Calculus and GREq.

Table 4.33 Univariable and multivariable comparison of odds ratio for the interaction of uGPA and GREq with taking calculus

	Univariable OR	**Multivariable Adj OR**
uGPA X Calculus	16.52 (95% CI:2.16 – 126.23)	8.25 (95% CI: 3.16 – 21.54)
GREq	10.49 (95% CI: 4.11 – 26.78)	9.59 (95% CI: 1.20 – 76.70)

This table demonstrates that GREq appears to have an independent effect (10.49 – 9.59), but uGPA and calculus clearly interact. A 2 X 2 analysis that does not include GREq (uGPA X Calculus) generates an OR that is different from the multivariable adjusted OR derived from a logistic regression analysis that did include calculus (16.52 – 8.25).

The next series of tables and figures demonstrates the nature of the interactive relationship between GREq and taking calculus for prediction of gGPA (Tables 4.34 to 4.36 and Figure 4.6).

Table 4.34 A student with a combination of a high GREq (≥ 141.5 [PR ≥ 12]) and took calculus for prediction of first-year gGPA ≥ 3.45

	First-year gGPA of ≥ 3.45	**First-year gGPA of < 3.45**
Both factors, GREq X Calculus (1)	65	5
≤ 1 factor, either GREq X Calculus (0)	29	35
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.69 (95% CI: 0.59 – 0.78)	Sp = 0.88 (95% CI: 0.74 – 0.95)	
OR = 15.69 (95% CI: 5.58 – 44.13)	RFS = 2.05 (95% CI: 1.68 – 2.51)	

A student who had both a high GREq (≥ 141.5 [PR ≥ 12]) and took calculus had 15.69 times greater odds to be successful in the GATP than the odds for someone who had one or none of these factors. The relative frequency of GATP success indicates the probability of a student being successful in the GATP who had both a high GREq and had taken calculus as an undergraduate is slightly more than twice that for a student who has only one or none of these factors.

Table 4.35 Stratified analysis of calculus history for association of GREq as predictor of gGPA

Calculus - Yes				
	Success	**Not successful**	**Total**	**Percentage**
High GREq	38	2	40	95%
Low GREq	3	1	4	75%
OR = 6.33				

Calculus - No				
	Success	**Not successful**	**Total**	Percentage
High GREq	47	16	63	75%
Low GREq	6	19	25	24%
OR = 9.30				

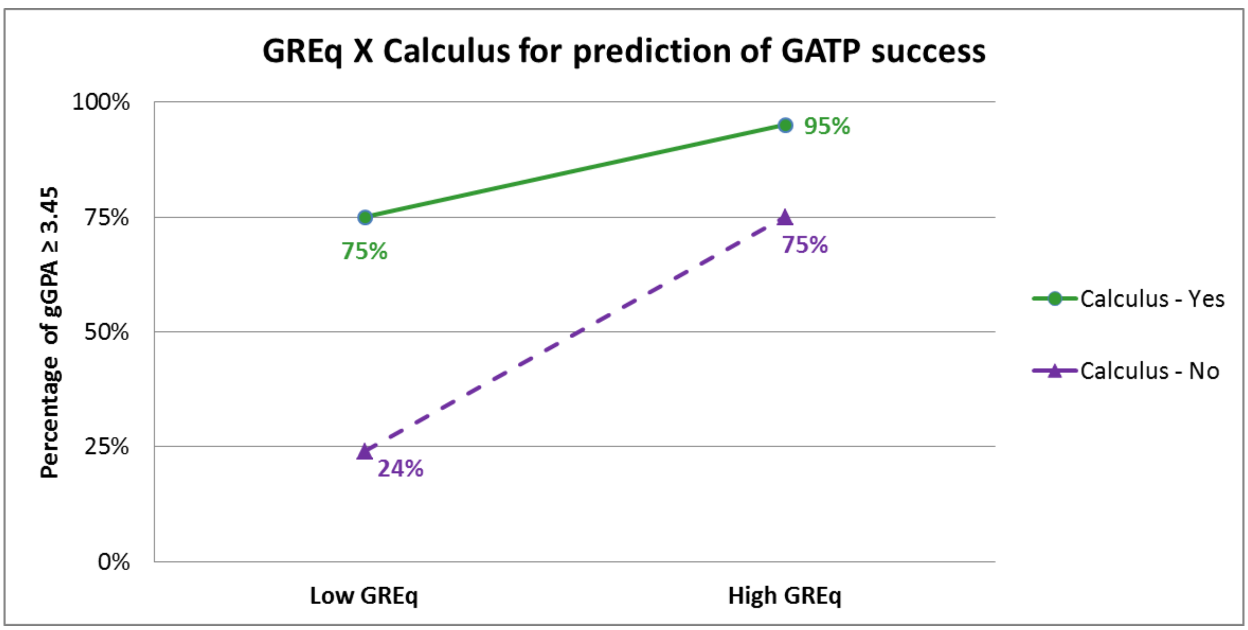

Figure 4.6 Calculus X GREq for prediction of GATP success

This possible interaction represents students who took calculus as an undergraduate tended to be successful regardless of their GREq score; 95% if they had a high GREq ($\geq$ 141.5 [PR $\geq$ 12]) versus 75% if they had a low GREq (< 141.5 [PR < 12]). If a candidate had a high GREq, but did not take calculus, 75% were successful, compared to only 24% who were successful if they had a low GREq and did not take calculus. The OR indicates that a student who had a high GREq and took calculus had 6.33 times greater odds to be successful in the GATP than the odds for someone who had a low GREq and took calculus. A student who had a high GREq and did not take calculus had 9.30 times greater odds to be successful in the GATP than the odds for someone who had a low GREq and did not take calculus.

Controlling for taking calculus, the relationship between GREq and GATP success (gGPA at the end of the first year $\geq$ 3.45) was examined using a Mantel-Haenszel analysis (Mantel-Haenszel OR_{est} = 8.97 [95% CI: 3.29 – 24.49]). There is a statistically significant association between GREq and GATP success (gGPA at the end of the first year $\geq$ 3.45) and high and low uGPA strata (taking calculus OR = 6.33 [95% CI: 0.44 – 91.71]); not taking calculus OR = 9.30 (95% CI: 3.15 – 44.12); Mantel-Haenszel $\chi^2(1) = 18.85$; $p < 0.001$). The null hypothesis for the Breslow-Day test assumes that the odds ratios for GREq by GATP success (gGPA at the end of the first year $\geq$ 3.45) is equivalent for taking versus not taking calculus categories. The Breslow-Day test for homogeneity found the odds ratios to not be significantly different from one another, Breslow-Day $\chi^2(1) = 0.070$; ($p = 0.791$). It should be noted due to the low cell counts (several < 5) make these results highly unstable (Hosmer & Lemeshow, 2000).

An examination of the univariable odds ratio and the multivariable adjusted odds ratio is shown in Table 4.36 for GREq X Calculus with uGPA.

Table 4.36 Univariable and multivariable comparison of odds ratio for the interaction of uGPA and GREq with taking calculus

	Univariable OR	**Multivariable Adj OR**
GREq X Calculus	13.57 (95% CI: 3.09 – 59.54)	14.90 (95% CI: 3.25 – 68.22)
uGPA	4.71 (95% CI: 2.17 – 10.23)	5.15 (95% CI: 2.21 – 12.01)

This table demonstrates that uGPA appears to have an independent effect (4.71 – 5.15), as do GREq and calculus. A 2 X 2 analysis that does not include uGPA (GREq X Calculus) generates an OR that is not very different from the multivariable adjusted OR derived from a logistic regression analysis that did include uGPA (13.57 – 14.90).

Three-way interaction

An examination of the three-way interaction between GREq (≥ 141.5 [PR ≥ 12]), took calculus and uGPA (≥ 3.18) was made. The 2 X 2 cross-tabulations table showing the results of this analysis is below (Table 4.37). Please note the upper right cell (All three factors and first-year gGPA of < 3.45) had zero (0) subjects in the cell. In order to compute the odds ratio, 0.5 was added to all cells (Hosmer & Lemeshow, 2000).

Table 4.37 A student with a combination of a high GREq score (≥ 141.5 [PR ≥ 12]), a high uGPA (≥ 3.18), and took calculus as an undergraduate for prediction of first-year gGPA ≥ 3.45

	First-year gGPA of ≥ 3.45	First-year gGPA of < 3.45
All three factors (GREq X Calculus X uGPA)	19.5	0.5
< 3 Factors	61.5	39.5

Fisher's Exact Test (one-sided) $p = 0.001$	
Sn = 0.24 (95% CI: 0.16 – 0.34)	Sp = 0.99 (95% CI: 0.89 – 1.00)
[a]OR = 25.05 (95% CI: 1.47 – 426.77)	RFS = 1.60 (95% CI: 1.31 – 1.96)

[a]OR calculated with 0.5 added to all cells

A student who had all three positive factors, (GREq ≥ 141.5 [PR ≥ 12]; took calculus; uGPA ≥ 3.18) had 25.05 times greater odds to be successful in the GATP than the odds for someone who had less than these three factors. The relative frequency of GATP success indicates the probability of a student being successful in the GATP who had a high GREq (≥ 141.5), took calculus, and had a high uGPA (≥ 3.18) is slightly more than one and half that for a student who does not have all three of these factors. Please note low cell counts (< 5) is cause for the fluctuations of the data and large confidence intervals; thus weakening the overall analysis and results (Hosmer & Lemeshow, 2000). Since the Fisher's Exact Test (one-sided) was statistically significant ($p = 0.001$) a graphic representation of the three-way interaction was created (Figure 4.7)

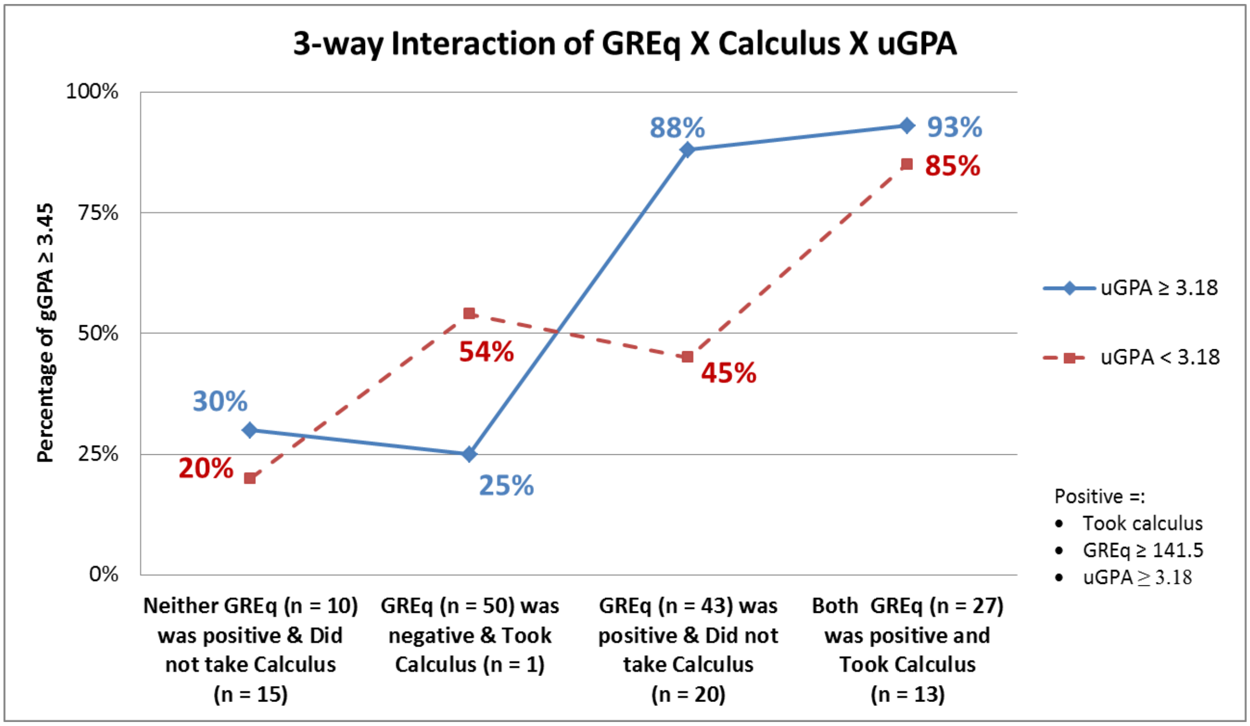

Figure 4.7 Three-way interaction of GREq X Calculus X uGPA for prediction of GATP success

The interaction indicates that students who were positive on the GREq (≥ 141.5) regardless of whether they took calculus and regardless of uGPA, had a high rate of success (uGPA ≥ 3.18, GREq ≥ 141.5 and took calculus = 93%; uGPA < 3.18, GREg ≥ 141.5 and did not take calculus = 88%). Those students who had a low uGPA, positive on the GREq, and took calculus had a high rate of success (uGPA < 3.18, GREg ≥ 141.5 and took calculus = 85%). Students who were negative on the GREq (≥ 141.5), took calculus, and had a high uGPA (< 3.18), were successful only 54% of the time. Students who were negative on the GREq (≥ 141.5), and took calculus, but had a low uGPA (< 3.18), only 25%were successful. Caution should be taken in interpreting this result as only one student took calculus in this category.

Regardless of uGPA, students having a low GREq score and not taking calculus were not very successful (uGPA $\geq$ 3.18 = 30%; uGPA < 3.18 = 20%).

Biderman's Formula Score

Interaction Effects

The second logistic regression analysis included Biderman's Formula Score. Potential interaction term included in this analysis was: Biderman's Score X uGPA; Calculus X uGPA; Biderman's Formula Score X Calculus; for prediction of GATP success (success = gGPA $\geq$ 3.45 at the end of the first year). Each set of interactions were examined in the same manner as the previous set of analyses.

The univariable odds ratio and multivariable adjusted odds ratio for each of the predictor variables is shown in Table 4.38.

Table 4.38 Comparison of odds ratios for predictor variables

	Univariable OR	**Multivariable Adj OR**
uGPA	4.71 (95% CI: 2.17 – 10.23)	2.55 (95% CI: 0.95 – 6.86)
Biderman's Score	16.94 (95% CI: 4.81 – 59.66)	8.34 (95% CI: 2.17 – 32.06)
Calculus	10.06 (95% CI: 2.90 – 34.86)	6.49 (95% CI: 1.67 – 25.23)

The existence of an interaction between the univariable odd ratios and the adjusted odds ratios is suggested by the differences between the univariable OR and the corresponding multivariable adjusted OR. The next series of tables and figures examine the relationship between Biderman's Score and uGPA (Tables 4.39 to 4.41 and Figure 4.8).

Table 4.39 A student with a combination of a high Biderman's Formula Score (≥ 458.45) and a high uGPA (≥ 3.18) for prediction of first-year gGPA ≥ 3.45

	First-year gGPA of ≥ 3.45	**First-year gGPA of < 3.45**
Both factors, Biderman X uGPA (1)	47	1
≤ 1 Factor, either Biderman X uGPA (0)	47	41
Fisher's Exact Test (one-sided) *p* < 0.001		
Sn = 0.50 (95% CI: 0.40 – 0.60)	Sp = 0.98 (95% CI: 0.88 – 1.00)	
OR = 41.00 (95% CI: 5.41 – 310.47)	RFS = 1.83 (95% CI: 1.67 – 2.51)	

A student who had a both high Biderman's Formula Score (≥ 458.45) and had a high uGPA (≥ 3.18) had 41.00 times greater odds to be successful in the GATP than the odds for someone who had only one or none of the factors. The relative frequency of GATP success indicates the probability of a student being successful in the GATP who had both a high Biderman's Formula Score and a high uGPA is 1.83 times greater probability of a student who had only one or none of these factors. Please note the cell count of "1" is cause to interpret these results with skepticism since it weakens the overall analysis and results in highly unstable odd ratios (Hosmer & Lemeshow, 2000).

Table 4.40 Stratified analysis of uGPA levels for association of Biderman's Formula Score as a predictor of gGPA

uGPA ≥ 3.18				
	Success	**Not successful**	**Total**	**Percentage**
High Biderman	47	1	48	**98%**
Low Biderman	21	14	35	**60%**
OR = 33.57				

uGPA < 3.18				
	Success	**Not successful**	**Total**	**Percentage**
High Biderman	7	2	9	**78%**
Low Biderman	19	25	44	**43%**
OR = 4.06				

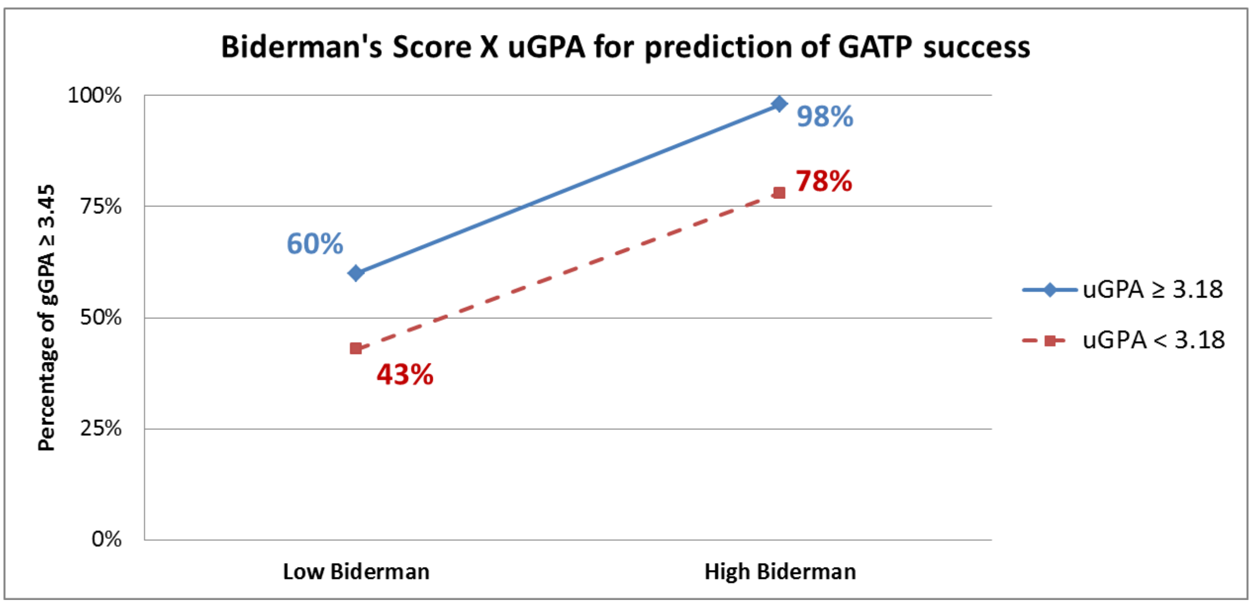

Figure 4.8 Biderman's Formula Score X uGPA for the prediction of GATP success

The possible interaction indicates that students who had a high Biderman's Formula Score ($\geq$ 458.5) were successful regardless of uGPA (uGPA $\geq$ 3.45 = 98%; uGPA < 3.45 = 78%). A student who a high uGPA ($\geq$ 3.45) and had a high Biderman's Formula Score ($\geq$ 458.5) had 33.57 times greater odds for success in the GATP than the odds for someone who had high uGPA and a low Biderman's Formula Score (< 458.5). A student who a low uGPA (< 3.45) and had a high Biderman's Formula Score ($\geq$ 458.5) had 4.06 times greater odds for success in the GATP than the odds for someone who had low uGPA and a low Biderman's Formula Score (< 458.5). A statistical anomaly demonstrates an interaction effect between Biderman's Formula Score and uGPA. Figure 4.8 demonstrates that it is not a true interaction effect, but the statistical interaction effect resolves the divergence of the 2 X 2 ORs and adjusted ORs.

Controlling for uGPA, the relationship between Biderman's Formula Score and GATP success (gGPA at the end of the first year $\geq$ 3.45) was examined using a Mantel-Haenszel analysis (Mantel-Haenszel OR_{est} = 11.58 [95% CI: 3.34 – 40.15]) and for homogeneity the Breslow-Day test. There is a statistically significant association between Biderman's Formula Score and GATP success (gGPA at the end of the first year $\geq$ 3.45) and high and low uGPA strata ($\geq$ 3.18 OR = 31.33 [95% CI: 3.86 – 254.08]; < 3.18 OR = 4.61 [95% CI: 0.86 – 24.73]); Mantel-Haenszel $\chi^2(1) = 11.577$; $p < 0.001$). The null hypothesis for the Breslow-Day test assumes that the odds ratios for Biderman's Formula Score by gGPA at the end of the first year is equivalent for uGPA categories. The Breslow-Day test for homogeneity found the odds ratios to not be significantly different from one another, Breslow-Day $\chi^2(1) = 2.158$; ($p = 0.142$). Please note the large confidence intervals are due to the low cell counts. Hosmer & Lemeshow (2000) suggest a minimum of five for each cell to have more reliable, valid, and stable model.

Table 4.41 Univariable and multivariable comparison of odds ratio for the interaction of Biderman's Formula Score and uGPA with taking calculus

	Univariable OR	**Multivariable Adj OR**
Biderman's X uGPA	41.00 (95% CI: 5.41 – 310.47)	37.58 (95% CI: 4.87 – 290.25)
Calculus	10.06 (95% CI: 2.90 – 34.86)	8.95 (95% CI: 2.43 – 32.92)

This tables demonstrates taking calculus has an independent effect (10.06 – 8.95), but there appears to be an interaction between Biderman's Formula Score and uGPA. A 2 X 2 analysis that does not include calculus (Biderman's Formula Score X uGPA) generates an OR that is different from the multivariable adjusted OR derived from a logistic regression analysis that did include uGPA (41.00 – 37-58).

The next series of tables and figures examine the relationship between taking calculus and uGPA (Tables 4.42 to 4.43 and Figure 4.9).

Table 4.42 A student with a combination of a high uGPA (≥ 3.18) and took calculus as an undergraduate for prediction of first-year gGPA ≥ 3.45

	First-year gGPA of ≥ 3.45	**First-year gGPA of < 3.45**
Both factors, uGPA X Calculus (1)	27	1
≤ 1 factor, either uGPA X Calculus (0)	67	41
Fisher's Exact Test (one-sided) *p* < 0.001		
Sn = 0.29 (95% CI: 0.21 – 0.39)	Sp = 0.97 (95% CI: 0.88 – 1.00)	
OR = 16.52 (95% CI: 2.16 – 126.23)	RFS = 1.54 (95% CI: 1.27 – 1.91)	

A student who had both a high uGPA (≥ 3.18) and had taken calculus as an undergraduate had 16.52 times greater odds to be successful in the GATP than the odds for someone who had only one or none of the factors. The relative frequency of GATP success indicates the probability of a student being successful in the GATP who had both a high uGPA and had taken calculus as an undergraduate is slightly more than one and half times the greater probability of a student who had only one or none of these factors. Please note the cell count of "1" is cause to interpret these results with skepticism since it weakens the overall analysis and results in highly unstable odd ratios (Hosmer & Lemeshow, 2000).

Table 4.43 Stratified analysis of uGPA levels for association of calculus history as a predictor of gGPA

uGPA ≥ 3.18				
	Success	**Not successful**	**Total**	**Percentage**
Calculus - Yes	27	1	28	96%
Calculus - No	41	14	55	75%
OR = 9.22				

uGPA < 3.18				
	Success	**Not successful**	**Total**	**Percentage**
Calculus - Yes	14	2	16	88%
Calculus - No	12	25	37	32%
OR = 14.58				

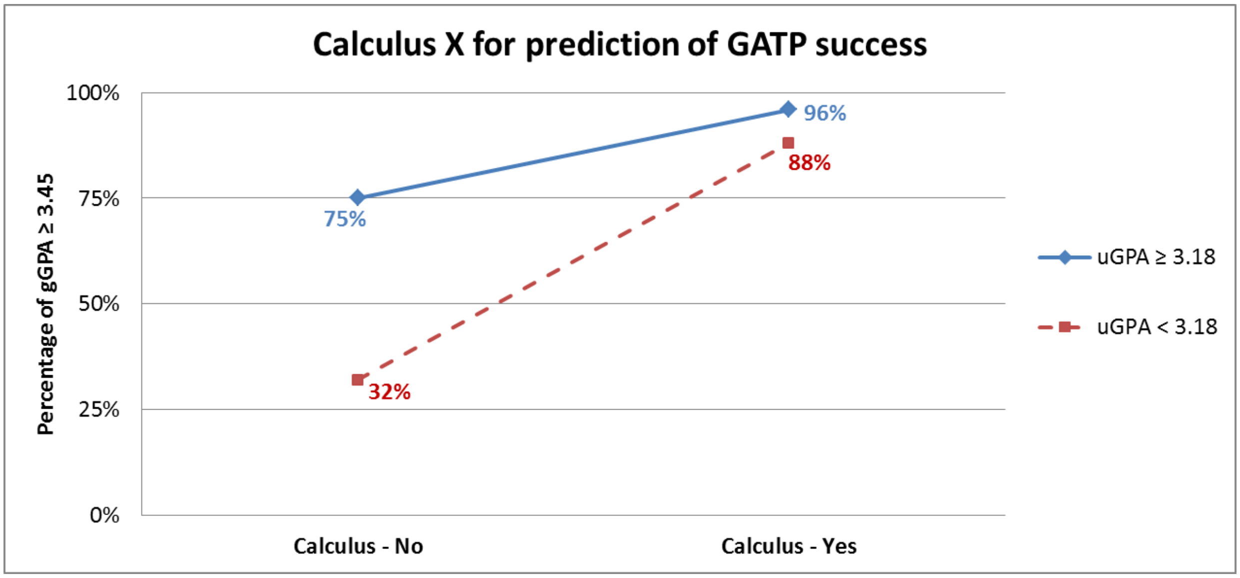

Figure 4.9 Calculus X uGPA for prediction of GATP success

The interaction indicates that students who took calculus had a high rate of success regardless of uGPA (uGPA $\geq$ 3.18 = 96%; uGPA < 3.18 = 88%). A student who took calculus and who had a high uGPA ($\geq$ 3.18) had 9.22 times greater odds for success in the GATP than the odds for someone with a high uGPA, who did not take calculus. Students who took calculus, but had a low uGPA (< 3.18) had 14.58 times greater odds for success in the GATP than the odds for someone who had a low uGPA and did not take calculus.

Controlling for uGPA, the relationship between taking calculus and GATP success (gGPA at the end of the first year $\geq$ 3.45) was examined using a Mantel-Haenszel analysis (Mantel-Haenszel OR_{est} = 11.79 [95% CI: 3.15 – 44.12]). There is a statistically significant association between taking calculus and GATP success (gGPA at the end of the first year $\geq$ 3.45) and high and low uGPA strata ($\geq$ 3.18 OR = 9.22 [95% CI: 1.15 – 74.25]; < 3.18 OR = 14.58 [95% CI: 2.85 – 74.71]); Mantel-Haenszel $\chi^2(1) = 16.76$; ($p < 0.001$). The null hypothesis for the Breslow-Day test assumes that the odds ratios for taking calculus by gGPA at the end of the first year is equivalent for uGPA categories. The Breslow-Day test for homogeneity found the odds ratios to be significantly different from one another, Breslow-Day $\chi^2(1) = 6.045$; ($p = 0.014$). Please note the large confidence intervals are due to the low cell counts. Hosmer & Lemeshow (2000) suggest a minimum of five for each cell to have more reliable, valid, and stable model.

An examination of the univariable odds ratio and the multivariable adjusted odds ratio is shown in Table 4.44 for uGPA X Calculus and Biderman's Formula Score.

Table 4.44 Univariable and multivariable comparison of odds ratios for the interaction of calculus and uGPA with Biderman's Formula Score

	Univariable OR	**Multivariable Adj OR**
Calculus X uGPA	16.52 (95% CI:2.16 – 126.23)	7.63 (95% CI:0.919 – 63.31)
Biderman's Formula Score	17.55 (95% CI: 5.06 – 60.86)	12.87 (95% CI:3.64 – 45.55)

This table demonstrates that Biderman's Formula Score appears to have an interaction effect (17.55 – 12.87), and calculus X uGPA also interact. A 2 X 2 analysis that does not include Biderman's Formula Score (Calculus X uGPA) generates an OR that is different from the multivariable adjusted OR derived from a logistic regression analysis that did include Biderman's Formula Score (16.52 – 7.63).

The next series of tables and figures examine the relationship between Biderman's Score and calculus (Tables 4.45 to 4.47 and Figure 4.10).

Table 4.45 A student with a combination of a high Biderman's Formula Score (≥ 458.5) and took calculus as an undergraduate for prediction of first-year gGPA ≥ 3.45

	First-year gGPA of ≥ 3.45	**First-year gGPA of < 3.45**
Both factors, Biderman's X Calculus (1)	28	1
≤ 1 factor, either Biderman's X Calculus (0)	66	41
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.30 (95% CI: 0.21 – 0.40)	Sp = 0.98 (95% CI: 0.88 – 1.00)	
OR = 17.39 (95% CI: 2.28 – 132.75)	RFS = 1.57 (95% CI: 1.27 – 2.92)	

A student who had both a high Biderman's Formula Score (≥ 458.5) and took calculus had 17.39 times greater odds to be successful in the GATP than the odds for someone who had only one or none of these factors. The relative frequency of GATP success indicates the probability of a student being successful in the GATP who had both a high Biderman's Formula Score (≥ 458.5) and took calculus was slightly more than one and half times greater probability of a student who had only one or none of these factors. Please note the cell count of "1" is cause to interpret these results with skepticism since it weakens the overall analysis and results in highly unstable odd ratios (Hosmer & Lemeshow, 2000).

Table 4.46 Stratified analysis of Biderman's Formula Score for levels of association of calculus history as a predictor of gGPA

Biderman ≥ 458.45				
	Success	**Not successful**	**Total**	**Percentage**
Calculus - Yes	28	1	29	**97%**
Calculus - No	26	2	28	**93%**
OR = 2.15				

Biderman < 458.45				
	Success	**Not successful**	**Total**	**Percentage**
Calculus - Yes	13	2	15	**87%**
Calculus - No	27	37	64	**42%**
OR = 8.10				

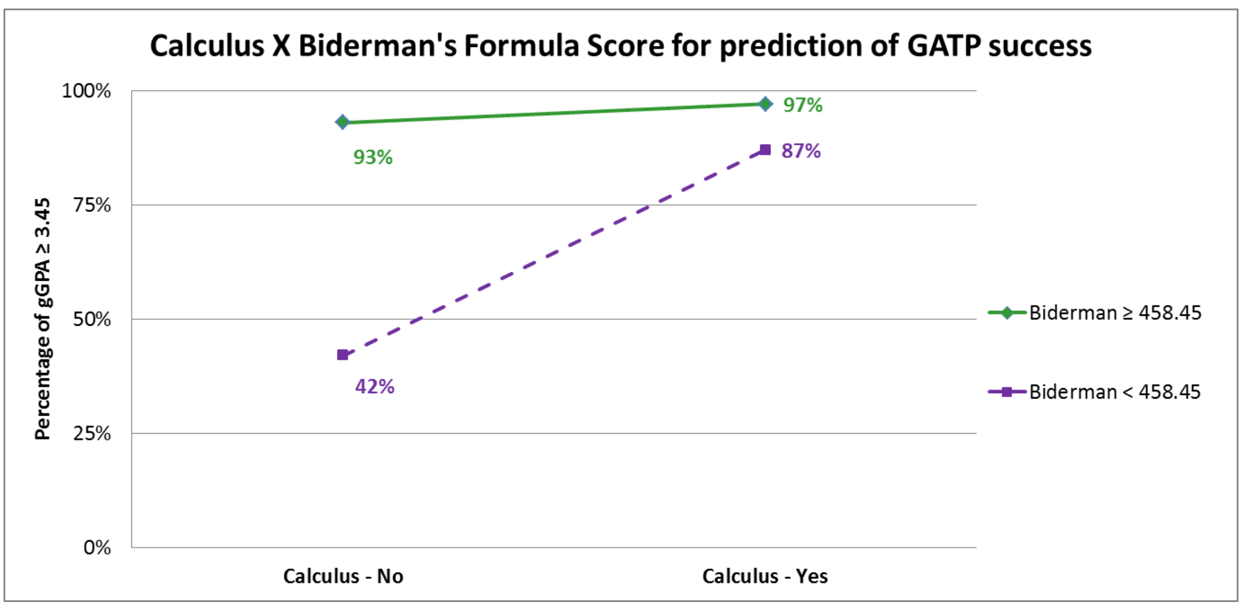

Figure 4.10 Calculus X Biderman's Formula Score for the prediction of GATP success

The interaction represents student who had a high Biderman's Formula Score (≥ 458.5) tended to be successful regardless of whether or not they took calculus; 97% if they had taken calculus versus 93% if they had not taken calculus. Students with a low Biderman's Formula Score (< 458.5) and took calculus, 87% were successful compared to only 42% who were successful if they had a low Biderman's Formula Score and did not take calculus. The OR indicates a student who had a high Biderman's Formula Score (≥ 458.5) and took calculus had 2.15 times greater odds to be successful in the GATP than the odds for someone who had a high Biderman's Formula Score and had not taken calculus. Students who had a low Biderman's Formula Score (< 458.5) and took calculus had 8.10 times great odds to be successful in the GATP than the odds for someone who had a low Biderman's Formula Score and had not taken calculus.

Controlling for Biderman's Formula Score (≥ 458.5), the relationship between taking calculus and GATP success (gGPA at the end of the first year ≥ 3.45) was examined using a Mantel-Haenszel analysis (Mantel-Haenszel $OR_{est} = 6.20$ [95% CI: 1.71 – 22.53]) and for homogeneity the Breslow-Day test. There is a statistically significant association taking calculus and GATP success (gGPA at the end of the first year ≥ 3.45) and high and low Biderman's Formula Score strata (≥ 458.5 OR = 2.154 [95% CI: 0.18 – 25.19]; < 458.5 OR = 8.10 [95% CI: 1.86 – 42.78]); Mantel-Haenszel $\chi^2(1) = 7.764$; ($p = 0.005$). The null hypothesis for the Breslow-Day test assumes that the odds ratios for taking calculus by gGPA at the end of the first year is equivalent for Biderman's Formula Score categories. The Breslow-Day test for homogeneity found the odds ratios to not be significantly different from one another, Breslow-Day $\chi^2(1) = 0.980$; ($p = 0.322$). Please note the large confidence intervals are due to the low cell counts. Hosmer & Lemeshow (2000) suggest a minimum of five for each cell to have more reliable, valid, and stable model.

Table 4.47 Univariable and multivariable comparison of odds ratio for the interaction of calculus and uGPA with Biderman's Formula Score

	Univariable OR	**Multivariable Adj OR**
Biderman's Score X Calculus	17.39 (95% CI: 2.28 – 132.75)	15.46 (95% CI: 1.98 – 121.02)
uGPA	4.71 (95% CI: 2.17 – 10.23)	4.31 (95% CI: 1.91 – 9.71)

This table demonstrates that uGPA appears to have an independent effect (4.71 – 4.31), and Biderman's Formula Score and calculus appear to have an interaction effect. A 2 X 2 analysis that does not include uGPA (Biderman's Formula Score X Calculus) generates an OR that is different from the multivariable adjusted OR derived from a logistic regression analysis that did include uGPA (17.39 – 15.46).

Three-way interaction

An examination of the three-way interaction between Biderman's Formula Score, taking calculus and uGPA was made. The 2 X 2 cross-tabulations table showing the results of this analysis is below (Table 4.48). Please note the upper right cell (All three factors & first-year gGPA of < 3.45) had zero (0) subjects in the cell. In order to computer the odds ratio, 0.5 was added to all cells (Hosmer & Lemeshow, 2000).

Table 4.48 A student with a combination of a high Biderman's Formula Score, a high uGPA, and took calculus as an undergraduate for prediction of first-year gGPA ≥ 3.45

	First-year gGPA of ≥ 3.45	**First-year gGPA of < 3.45**
All three factors (BID X Calculus X uGPA)	15.5	0.5
< 3 Factors	65.5	39.5
Fisher's Exact Test (one-sided) $p = 0.002$		
Sn = 0.19 (95% CI: 0.12 – 0.29)	Sp = 0.99 (95% CI: 0.89 – 1.00)	
[a]OR = 18.69 (95% CI: 1.09 – 321.16)	RFS = 1.55 (95% CI: 1.27 – 1.90)	

[a]OR calculated with 0.5 added to all cells

A student who had all three positive factors, (Biderman's Formula Score ≥ 458.5; took calculus; uGPA ≥ 3.18) had 18.69 times greater odds to be successful in the GATP than the odds for someone who had less than these three factors. The relative frequency of GATP success indicates the probability of a student being successful in the GATP who has a high Biderman's Formula Score (≥ 458.5), has taken calculus, and has a high uGPA (≥ 3.18) is slightly more than one and half times greater probability of a student who does not have all three of these factors. Please note low cell counts (< 5) is cause for the fluctuations of the data and large confidence intervals; thus weakening the overall analysis and results (Hosmer & Lemeshow, 2000). Since the Fisher's Exact Test (one-sided) was statistically significant ($p = 0.002$) a graphic representation of the three-way interaction was created (Figure 4.11)

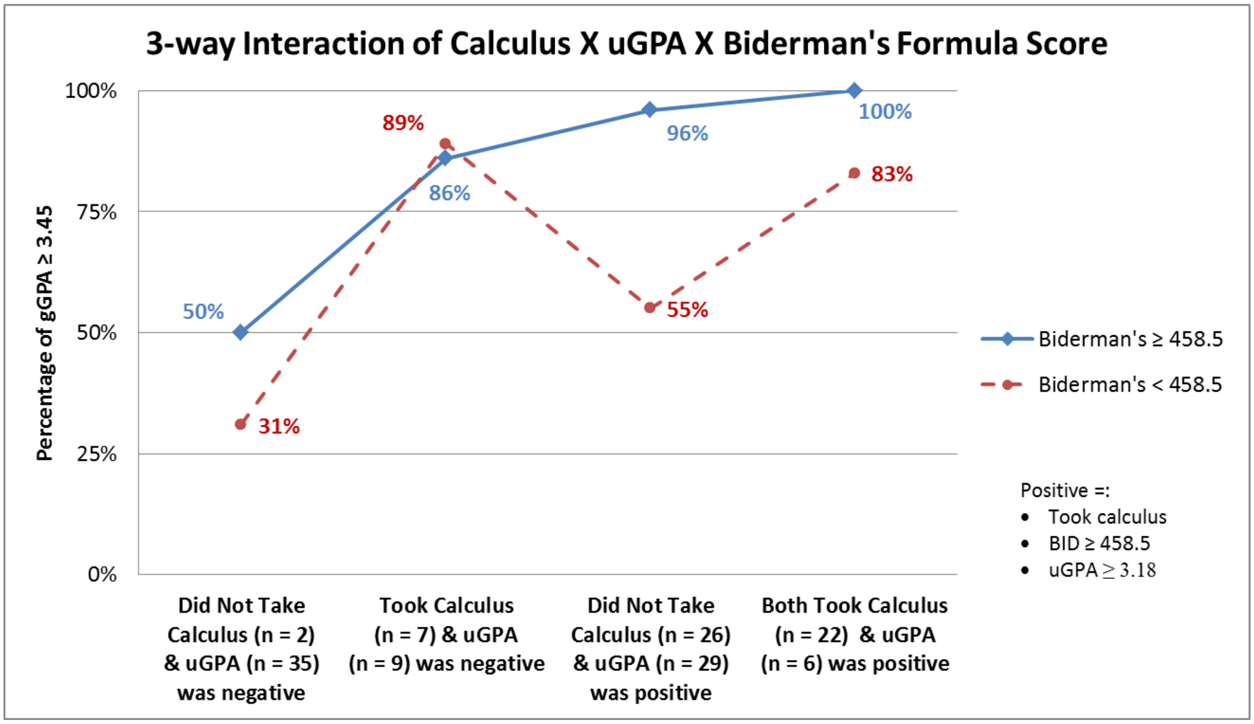

Figure 4.11 Three-way interaction of taking Calculus X uGPA X Biderman's Formula Score for prediction of gGPA ≥ 3.45

The interaction indicates that students who took calculus and who were positive for uGPA (≥ 3.18) and had a high Biderman's Formula Score (≥ 458.5) were all successful. Students who had a high Biderman's Formula Score regardless of whether or not they took calculus or what their uGPA was tended to be successful (Biderman's Formula Score ≥ 458.5, did not take calculus, uGPA < 3.18 = 96%; Biderman's Formula Score ≥ 458.5, took calculus, uGPA ≥ 3.18 = 86%). Students who took calculus, regardless of their uGPA, but had a low Biderman's Formula Score also tended to be successful (took calculus, uGPA ≥ 3.18, Biderman's Formula Score < 458.5 = 89%; took calculus, uGPA < 3.18, Biderman's Formula Score < 458.5 = 83%). Only half of the students who had a high Biderman's Formula Score, did not take calculus and had a low uGPA were successful (50%). Students with a low Biderman's Formula Score, did not take calculus and had a low uGPA were successful only 31% of the time.

Prediction Model

To create a final prediction model, the sum of the number of positive variables was used as a single variable with four levels (i.e., 0, 1, 2, or 3). Receiver operating characteristic analysis was used to identify the optimum number of positive factors for prediction of first-year gGPA. The results of ROC analyses for two different three-factor models are depicted in Figure 4.12 and Table 4.49 and Figure 4.13 and Table 4.51.

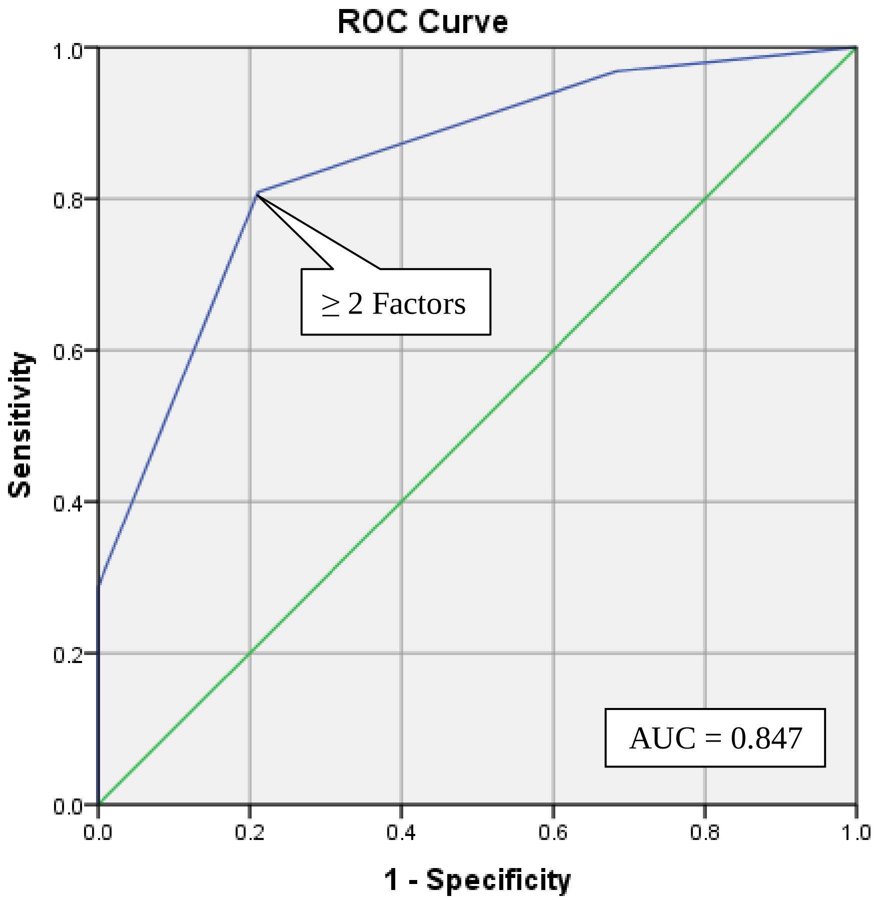

Figure 4.12 ROC curve with identification of the optimum cut-point for the number of positive factors (out of 3 factors) for prediction of success in the GATP as indicated by gGPA at the end of the first year ≥ 3.45 (includes GRE scores)

Table 4.49 Number of positive factors (out of three), for prediction of success in the GATP as indicated by gGPA at the end of the first year ≥ 3.45 (includes GRE scores)

	First-year gGPA of ≥ 3.45	**First-year gGPA of < 3.45**
≥ 2 Factors	76	8
< 2 Factors	18	34
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.81 (95% CI: 0.72 – 0.88)	Sp = 0.81 (95% CI: 0.67 – 0.90)	
Youden's Index = 0.598		
OR = 17.94 (95% CI: 7.11 – 45.29)	RFS = 2.61 (95% CI: 2.13 – 3.20)	

This prediction model found three positive factors: uGPA ≥ 3.18, GREq ≥ 141.5 (PR ≥ 12.0), and the student took calculus. A cut-point of two or more factors was found for optimum balance of Sn and Sp. A student in the GATP who had any combination of two or more of the three factors had 17.94 times greater odds of being successful in the GATP than the odds for someone who had less than two of the three factors. The relative frequency of GATP success indicates the probability of a student being successful in the GATP with any two or more of the three factors was two and half times the probability of a student with less than two factors. The success rate (gGPA ≥ 3.45) for a given number of positive factors is presented in Table 4.50.

Table 4.50 Specific number of factors for a three factor model for prediction of first-year gGPA ≥ 3.45

Number of Positive Factors	Success in the GATP gGPA ≥ 3.45	gGPA < 3.45	Total	Percentage	Percentage above/ below cut point
0	3	16	19	15.79%	18/52 = 34.62%
1	15	18	33	45.45%	
2	49	9	57	85.96%	76/84 = 90.48%
3	27	0	27	100.00%	
Total	94	42	136	71.21%	

Students with two or more positive factors demonstrated a 90.48% success rate in the GATP, whereas only 34.62% of the students with less than two factors were deemed successful. Overall, regardless of the number of factors, 71.21% of all students were “successful” with a first-year gGPA ≥ 3.45 indicating the selection committee had made the correct assessment for a large proportion of the students admitted to the program.

Information related to another alternative three-factor prediction model, are shown in Figure 4.13 and Table 4.51.

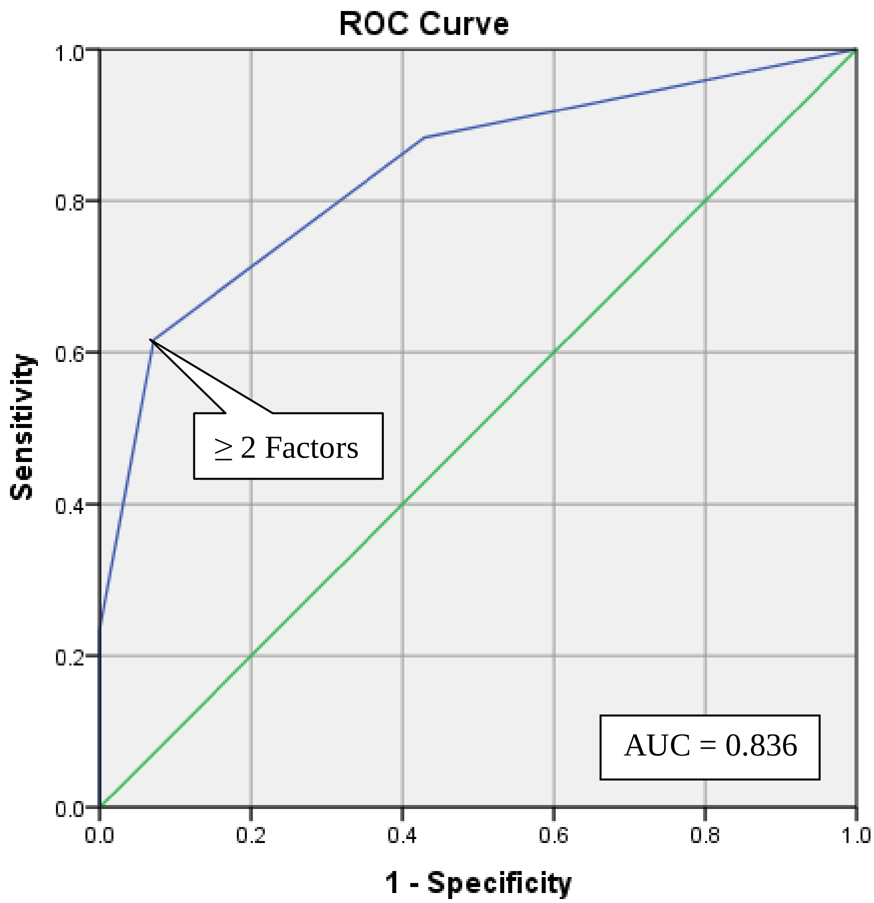

Figure 4.13 ROC curve with identification of the optimum cut-point for the number of positive factors for prediction of success in the GATP as indicated by gGPA at the end of the first year ≥ 3.45 (includes Biderman's Formula Score)

Table 4.51 Number of factors for prediction of success in the GATP as indicated by gGPA at the end of the first year ≥ 3.45 (includes Biderman's Formula Score)

	First-year gGPA of ≥ 3.45	**First-year gGPA of < 3.45**
≥ 2 Factors	58	3
< 2 Factors	36	39
Fisher's Exact Test (one-sided) *p* < 0.001		
Sn = 0.62 (95% CI: 0.52 – 0.71)	Sp = 0.93 (95% CI: 0.81 – 0.98)	
Youden's Index = 0.546		
OR = 20.94 (95% CI: 6.03 – 72.79)	RFS = 1.98 (95% CI: 1.62 – 2.43)	

The alternative three-factor prediction model for determining success in the GATP included Biderman's Formula Score ≥ 458.45, uGPA ≥ 3.18, and took calculus. A student in the GATP who had any combination of two or more of the three factors had 20.94 times greater odds of being successful in the GATP than the odds of someone who had less than two of the three factors. The relative frequency of GATP success indicates the probability of a student being successful in the GATP who had two or more of the three factors were almost twice that of a student with less than two of these factors. The success rate (gGPA ≥ 3.45) for a given number of positive factors is presented in Table 4.52.

Table 4.52 Specific number of factors for a three factor model for prediction of first-year gGPA ≥ 3.45

	Success in the GATP				
Number of Positive Factors	**gGPA ≥ 3.45**	**gGPA < 3.45**	**Total**	**Percentage**	**Percentage above/ below cut point**
0	11	24	35	31.43%	36/75 = 48.0%
1	25	15	40	62.50%	
2	36	3	39	92.31%	58/61 = 95.08%
3	22	0	22	100.0%	
Total	94	42	136	69.12%	

Students with two or more positive factors demonstrated a 93.94% success rate in the GATP, whereas only 48.0% of the students with less than two factors were deemed successful. Overall, regardless of the number of factors, 69.12% of all students were "successful" with a first-year gGPA ≥ 3.45 indicating the selection committee had made the correct assessment on selecting students to be a part of the GATP under 70% of the time.

Final Assessment

This project began in an effort to try to identify predictors for success in a GATP and predict success on the BOC exam. The data gathered came from one specific GATP. A very strong predictor of BOC success was a gGPA at the end of the first-year of 3.45 (OR = 8.30, Table 4.2). It is not likely, nor reasonable to assume, all GATPs will have a cut-point of 3.45 for gGPA. In order for these results to have real utility in the athletic training profession two final prediction models were produced. All of the previously used predictor variables, except for gGPA, were entered into another logistic regression. The results of the logistic regression analyses, ROC analyses, and 2 X 2 cross-tabulation tables are presented.

All of the previous dichotomized predictors were entered into the logistic regression analysis with "first-attempt pass – Yes or No, on the BOC exam" as the outcome variable. The predictor variables entered into the logistic regression were: advanced math and science courses ≥ 3, GREv ≥ 145.5, GREq ≥ 143.5, GREwr ≥ 3.25, Physics – Yes or No, and Calculus – Yes or No. The results of the logistic regression analysis are displayed in Table 4.53.

Table 4.53 Logistic regression analysis results including all potential predictors of first-attempt BOC exam success

			95% C.I.	
		Adj. OR	**Lower**	**Upper**
Step 1	Advanced math and science courses ≥ 3	1.927	0.449	8.268
	GREv ≥ 145.5 (PR ≥ 26)	2.682	0.820	8.769
	GREq ≥ 143.5 (PR ≥ 16.5)	6.272	1.783	22.059
	GREwr ≥ 3.25 (PR ≥ 24.5)	2.542	0.663	9.753
	Physics – Yes or No	0.858	0.192	3.842
	Calculus – Yes or No	0.367	0.073	1.835
	Constant	0.417		
Step 2	Advanced math and science courses ≥ 3	1.796	0.497	6.486
	GREv ≥ 145.5 (PR ≥ 26)	2.664	0.815	8.704
	GREq ≥ 143.5 (PR ≥ 16.5)	6.118	1.782	21.007
	GREwr ≥ 3.25 (PR ≥ 24.5)	2.529	0.659	9.711
	Calculus – Yes or No	0.350	0.075	1.642
	Constant	0.412		
Step 3	GREv ≥ 145.5 (PR ≥ 26)	2.890	.898	9.297
	GREq ≥ 143.5 (PR ≥ 16.5)	5.911	1.747	20.003
	GREwr ≥ 3.25 (PR ≥ 24.5)	2.245	0.614	8.203
	Calculus – Yes or No	0.480	0.123	1.873
	Constant	0.513		
Step 4	GREv ≥ 145.5 (PR ≥ 26)	2.700	.866	8.416
	GREq ≥ 143.5 (PR ≥ 16.5)	4.857	1.579	14.942
	GREwr ≥ 3.25 (PR ≥ 24.5)	2.194	0.600	8.024
	Constant	0.495		
Step 5	GREv ≥ 145.5 (PR ≥ 26)	3.292	1.123	9.655
	GREq ≥ 143.5 (PR ≥ 16.5)	5.334	1.767	16.102
	Constant	0.780		

This model produced five steps, in which two potential steps can be considered for the final prediction model, Step 4 with three predictors (GREv ≥ 145.5 [PR ≥ 26], GREq ≥ 143.5 [PR ≥ 16.5], and GREwr ≥ 3.25 [PR ≥ 24.5]) and Step 5 with two predictors (GREv ≥ 145.5 [PR ≥ 26] and GREq ≥ 143.5 [PR ≥ 16.5]). The Nagelkerke R^2 is 0.290 at Step 4 and 0.273 at Step 5.

To help determine which model was the better choice, ROC analyses were performed for each of the final two steps of the logistic regression (Figures 4.14 and 4.15). This was followed by 2 X 2 cross-tabulation analysis for each step (Tables 4.54 and 4.55)

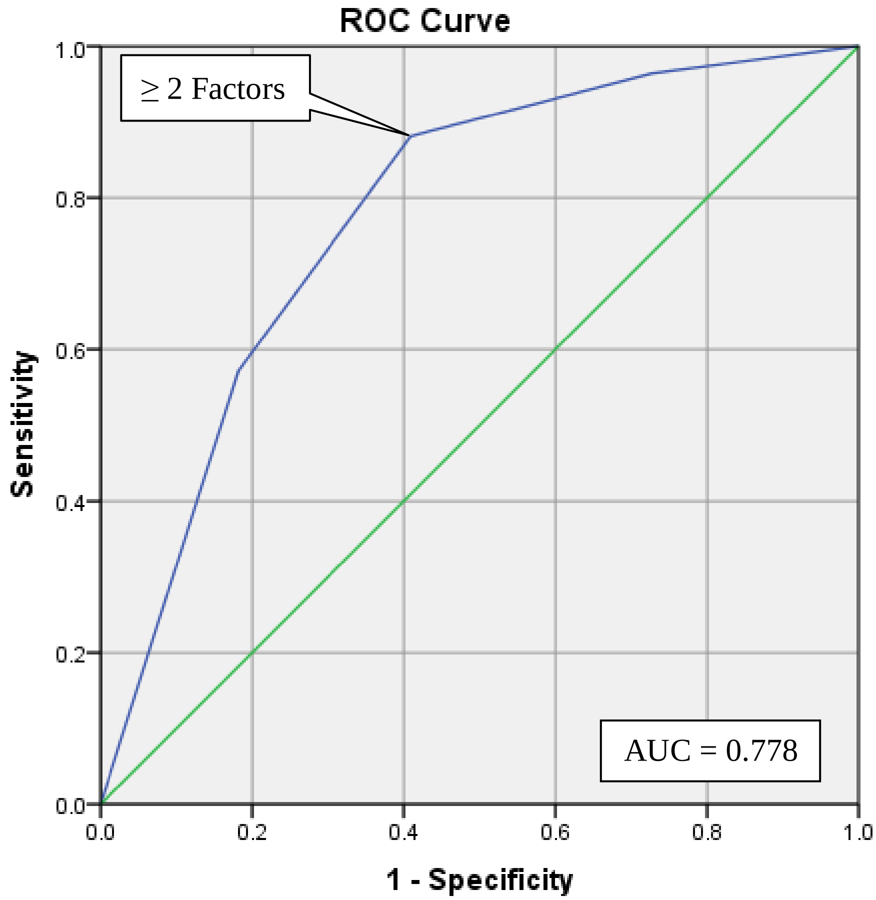

Figure 4.14 ROC curve with identification of the optimum cut-point for the number of positive factors (including GREv, GREq and GREwr scores) for prediction of first-attempt BOC exam success

Table 4.54 Number of factors (including GREv, GREq and GREwr scores) for prediction of first-attempt BOC exam success

	First-attempt Pass on the BOC exam	
	Yes	**No**
≥ 2 Factors	74	9
< 2 Factors	10	13
Fisher's Exact Test (one-sided) *p* < 0.001		
Sn = 0.88 (95% CI: 0.79 – 0.93)	Sp = 0.59 (95% CI: 0.39 – 0.77)	
Youden's Index = 0.472		
OR = 10.69 (95% CI: 3.64 – 31.36)	RFS = 2.05 (95% CI: 1.67 – 2.51)	

For the three-factor model a GATP student who had ≥ 2 positive factors, (GREv ≥ 145.5 (PR ≥ 26), GREq ≥ 143.5 [PR ≥ 16.5], GREwr ≥ 3.25 [PR ≥ 24.5]), had 10.69 times greater odds of first-attempt BOC exam success than the odds for someone who had less than two of the three factors. The relative frequency of success indicates the probability of a student passing the BOC exam on the first-attempt with any two of the three factors is slightly greater than twice the probability of a student with one or none of the three positive factors.

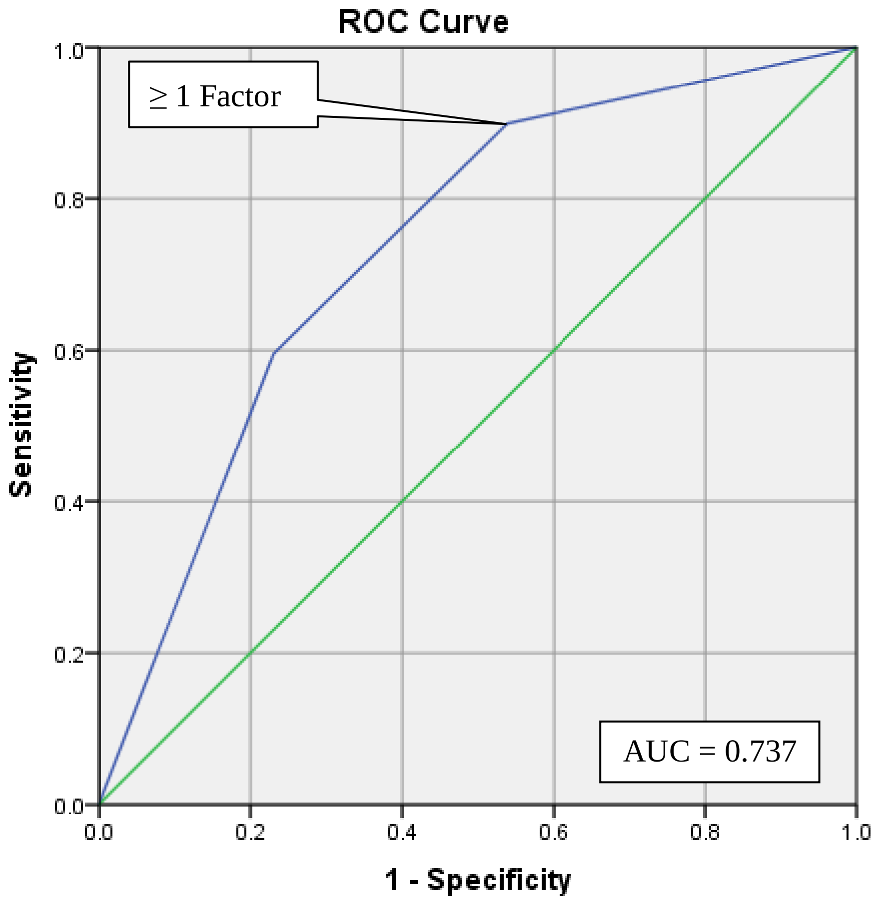

Figure 4.15 ROC curve with identification of the optimum cut-point for the number of positive factors (including GREv and GREq scores) for prediction of first-attempt BOC exam success

Table 4.55 Number of factors (including GREv and GREq scores) for prediction of first-attempt BOC exam success

	First-attempt Pass on the BOC exam	
	Yes	**No**
≥ 1 Factor	80	14
No Factors	9	12
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.90 (95% CI: 0.82 – 0.95)	Sp = 0.46 (95% CI: 0.29 – 0.65)	
Youden's Index = 0.361		
OR = 7.62 (95% CI: 2.71 – 21.43)	RFS = 1.99 (95% CI: 1.62 – 2.43)	

For the two-factor model a GATP student who had either one of the two positive factors, (GREv ≥ 145.5 (PR ≥ 26) or GREq ≥ 143.5 [PR ≥ 16.5]), had 7.62 times greater odds of first-attempt BOC exam success than the odds for someone who had none of the two factors. The relative frequency of GATP success indicates the probability of a student being successful in the GATP with any one of the two factors is about twice the probability of a student with none of the positive factors.

The process was repeated entering all of the dichotomized predictors were entered into the logistic regression analysis with "first-attempt pass – Yes or No, on the BOC exam" as the outcome variable. The predictor variables entered into the logistic regression were: advanced math and science courses ≥ 3, Biderman's Formula Score ≥ 420.5, Physics – Yes or No, and Calculus – Yes or No. The results of the logistic regression analysis are displayed in Table 4.56.

Table 4.56 Logistic regression analysis results including all potential predictors (including Biderman's Formula Score) of first-attempt BOC exam success

			95% C.I.	
		Adj. OR	**Lower**	**Upper**
Step 1	Advanced math and science courses ≥ 3	1.264	0.347	4.610
	Biderman's Formula Score ≥ 420.5	4.671	1.647	13.243
	Physics – Yes or No	1.858	0.491	7.032
	Calculus – Yes or No	0.591	0.141	2.468
	Constant	1.318		
Step 2	Biderman's Formula Score ≥ 420.5	4.670	1.650	13.220
	Physics – Yes or No	2.081	0.636	6.801
	Calculus – Yes or No	0.623	0.154	2.519
	Constant	1.358		
Step 3	Biderman's Formula Score ≥ 420.5	4.396	1.589	12.164
	Physics – Yes or No	1.703	0.626	4.634
	Constant	1.368		
Step 4	Biderman's Formula Score ≥ 420.5	4.615	1.680	12.679
	Constant	1.733		

This model produced four steps with the final step having only one variable remaining, Biderman's Formula Score ≥ 420.5. The Nagelkerke R^2 is 0.136 at Step 4. There was no reason for ROC analysis with only one predictor variable remaining in the model. The following 2 X 2 cross-tabulation analysis was performed (Table 4.57).

Table 4.57 Number of factors (including Biderman's Formula Score) for prediction of first-attempt BOC exam success

	First-attempt Pass on the BOC exam	
	Yes	**No**
Biderman's Formula Score ≥ 420.5	58	7
Biderman's Formula Score < 420.5	26	15
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.69 (95% CI: 0.59 – 0.78)	Sp = 0.68 (95% CI: 0.47 – 0.84)	
OR = 4.78 (95% CI: 1.74 – 13.12)	RFS = 1.41 (95% CI: 1.15 – 1.73)	

For the Biderman's Formula Score model a GATP student who had a Biderman's Formula Score of ≥ 420.5 had 4.78 times greater odds of first-attempt BOC exam success than the odds for someone who had a Biderman's Formula Score of < 420.5.

Summary of Chapter

Chapter IV presented the results of this study. There are two interrelated purposes, both of which pertained to the process of admitting students to a graduate professional program. The first component of this study involved the development of a prediction model to identify factors associated with eligibility and first-attempt success on the Board of Certification (BOC) examination for students who have completed a professional (entry-level) graduate athletic training program (GATP). The analyses produced two prediction models. The first model had three predictors, gGPA at the end of the first year ≥ 3.45, GREv ≥ 145.5, and GREq ≥ 143.5. A GATP student, who had ≥ 2 positive factors, had 6.31 times greater odds of first-attempt BOC exam success than the odds for someone who had none or only one of the three factors. The relative frequency of success indicates the probability of a student passing the BOC exam on the first-attempt with any two or more of these factors is slightly more than one and half times the probability of a student who has less than two of these factors.

The second model had two predictors, gGPA at the end of the first year ≥ 3.45, and Biderman's Formula Score ≥ 420.5. A GATP student who had at least one positive factor, had 10.69 times greater odds of BOC exam success on the first-attempt than the odds for someone who had neither of the two factors. The relative frequency of GATP success indicates the probability of a student being successful in the GATP with one or more factors is slightly greater than twice the probability of a student with none of the positive factors.

The second component utilized results from the first analysis to identify program applicant characteristics that were most likely to predict both academic success within the graduate professional program and subsequent success on the BOC exam. This also produced two prediction models. The first model produced three predictors; uGPA ≥ 3.18, GREq ≥ 141.5, and having taken calculus as an undergraduate. A student in the GATP who had any combination of two or more of the three factors had 17.94 times greater odds of being successful in the GATP than the odds for someone who had less than two of the three factors. The relative frequency of GATP success indicates the probability of a student being successful in the GATP with any two or more of the three factors was twice the probability of a student with less than two factors.

The second model also produced three predictors; uGPA ≥ 3.18, Biderman's Formula Score ≥ 458.45, and took calculus as an undergraduate. A student in the GATP who had any combination of two or more of the three factors had 20.94 times greater odds of being successful in the GATP than the odds of someone who had less than two of the three factors. The relative frequency of GATP success indicates the probability of a student being successful in the GATP who had two or more of the three factors were almost twice that of a student with less than two of these factors.

Since the data gathered for this study came from one specific GATP and gGPA was one of the strongest predictors, a subsequent analysis was performed. The logistic regression was repeated with all of the final set of predictors except for gGPA. Two prediction models were produced. The first had three predictors: GREv $\geq$ 145.5, GREq $\geq$ 143.5, and GREwr $\geq$ 3.25. A student who had any combination of two or more of the three positive factors had 10.69 times greater odds of first-attempt BOC exam success than the odds for someone who had less than two of the three factors. The relative frequency of success indicates the probability of a student passing the BOC exam on the first-attempt with any two of the three factors is slightly greater than twice the probability of a student with one or none of the three positive factors.

The second model produced only one predictor, Biderman's Formula Score, $\geq$ 420.5. A GATP student who had a Biderman's Formula Score of $\geq$ 420.5 had 4.78 times greater odds of first-attempt BOC exam success than the odds for someone who had a Biderman's Formula Score of $<$ 420.5.

CHAPTER V

DISCUSSION AND CONCLUSION

This final chapter of the dissertation will restate the research questions and review the major methods which were used. This chapter will summarize the results and discuss the implications of those results along with addressing potential future study.

This study had two interrelated hypotheses, both of which pertained to the process of admitting students to a professional graduate athletic training program. The first component of this study involved the development of a prediction model to identify factors associated with eligibility and first-attempt success on the Board of Certification (Board of Certification) examination for students who have been enrolled in a professional (entry-level) graduate athletic training program (GATP). The second component utilized the results of the first analysis to identify program applicant characteristics that are most likely to predict both academic success in the graduate professional program and subsequent success on the Board of Certification exam. The results of this study lead us to accept both of the experimental hypotheses and reject both null hypotheses.

In Chapter II, we reported that nine previous studies had been performed in an attempt to predict first-attempt success on the Board of Certification exam; however, none of the studies were successful in identifying potential predictors of success on the Board of Certification exam. The commonalities of those nine studies are they involved students from undergraduate athletic training education programs, and each of them used frequentist statistics to analyze their data.

Additionally, several educators for several medical professions have attempted to identify predictors of the most qualified (i.e., likely to succeed) applicants to their professional programs. All of the studies identified used frequentist statistics to analyze their data.

For this study, we chose to use Bayesian philosophy to create prediction models for success on the Board of Certification exam and to identify characteristics of those candidates who are likely to be successful in a graduate athletic training program. In order to accomplish this, we identified all potential predictors of success, then performed univariable analyses using receiver operating characteristic (ROC) analyses and 2 X 2 cross-tabulation calculations to narrow the selection of predictors. An examination of multicollinearity (or the degree of possible overlap between the variables) was done for the continuous and multi-level discrete variables before repeating the process for dichotomous variables. The remaining predictors were then entered into a logistic regression to identify the strongest combination of variables. For both the prediction of first-attempt success on the Board of Certification exam and success in the graduate athletic training program, two different prediction models were created. The remaining predictors were finally examined for their degree of interaction or independence.

To predict first-attempt success on the Board of Certification exam, the three-factor model included a graduate grade point average, Graduate Record Exam (GRE) verbal score, and Graduate Record Exam quantitative score. Any student with a combination of any two of these three predictors or all three of the predictors has over six times greater chance of passing the Board of Certification exam on their first-attempt than someone who has less than two of the predictors. This is known as the odds ratio. Another way of looking at these data is students with two or more of the three predictors are over one and half times more likely to pass the

Board of Certification exam on their first-attempt than students with less than two of the predictors. This is known as the relative frequency of success.

An alternative model for predicting first-attempt success on the Board of Certification exam had only two predictors, graduate grade point average and a Biderman's Formula Score. If a student had at least one of these two predictors, then they have over ten and half times greater chance of passing the Board of Certification exam than someone who had neither of the predictors. Stated another way, if a student has at least one of the two predictors, then he or she is twice as likely to pass the Board of Certification exam on their first-attempt compared to someone who did not have either of the predictors.

Graduate Athletic Training Program success – GRE prediction model explained

Success in the graduate athletic training program was defined as having a graduate grade point average at the end of the first-year of 3.45 or above. To predict success in the graduate athletic training program two models were created. The first model included three predictors comprising the student's undergraduate grade point average, Graduate Record Exam quantitative score, and that the student took calculus as an undergraduate. The receiver operating characteristic analysis demonstrates that any combination of two or more of the predictors identifies the cut-point (Figure 5.1). The odds ratio generated from the 2 X 2 cross tabulations table found any student with a combination of any two of these three predictors or all three of the predictors has almost eighteen times greater odds of being successful in a graduate athletic training program compared to a student who has either one or none of the predictors. Stated another way, a student with any combination of two or all three of the predictors are more than twice as likely to be successful in the graduate athletic training program compared to a student who does not have one or none of the predictors. This is known as the relative frequency of success.

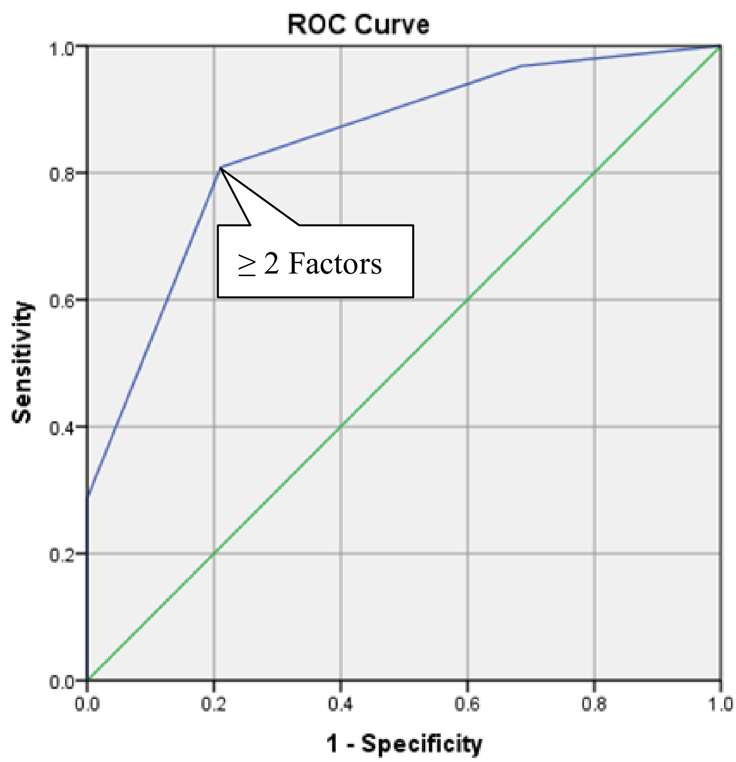

Figure 5.1 Receiver operating characteristic curve with identification of the optimum cut-point for the number of positive factors for prediction of success in the graduate athletic training program as indicated by graduate grade point average at the end of the first year ≥ 3.45 (includes GRE scores)

Although the relationship of having any combination of two of the three predictors is quite robust, it does not explain which combination of predictors is strongest. A series of analyses found students who had a high undergraduate grade point average and a high quantitative score on the Graduate Record Exam led to the greatest percentage of successful students in the graduate athletic training program. When adding the third predictor to the analysis, it is best to have a high undergraduate grade point average, a high GRE quantitative score, and to have taken calculus. Students who fit this profile were almost always successful in the graduate athletic training program. However, if students' with an undergraduate grade point average that was not as high, but they still had a high GRE quantitative score, and had taken calculus were still very successful (Figure 5.2).

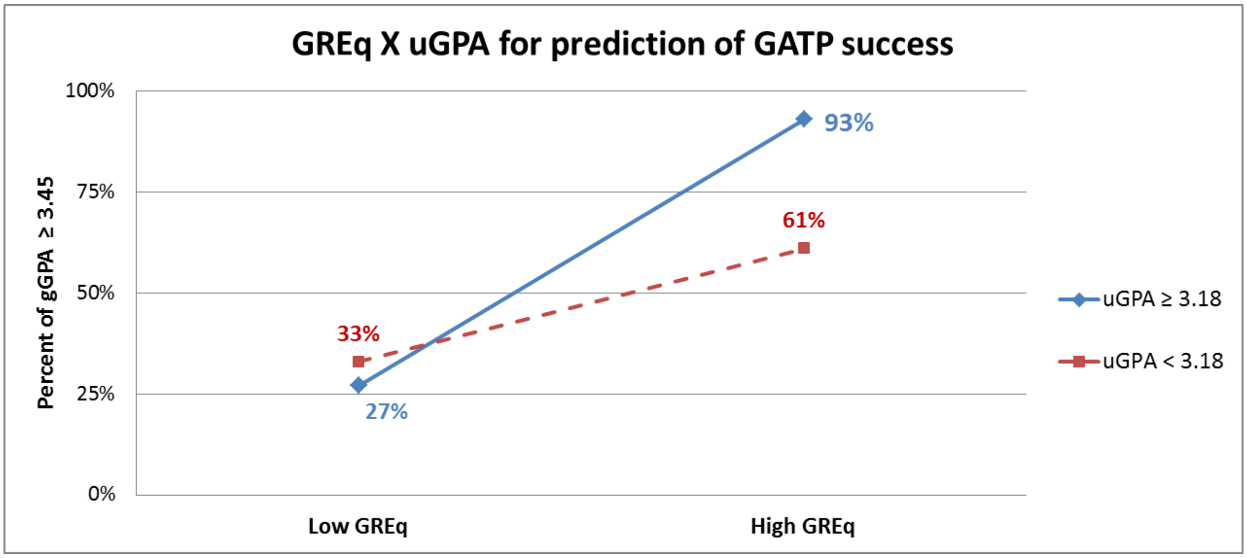

Figure 5.2 Interaction of GRE quantitative score (GREq) X undergraduate grade point average (uGPA) for prediction of graduate athletic training program success (graduate GPA at the end of the first year ≥ 3.45)

Graduate Athletic Training Program success – Biderman's Formula Score prediction model explained

Biderman's Formula Score was borrowed from the University of Tennessee at Chattanooga's Psychology Department's graduate application criteria. They did not explain or quantify how much more successful students were who had achieved a score of 480 or above over students who had a score below 480. Biderman's Formula Score involves a calculation of one's undergraduate grade point average times 100, plus the sum of the percentile ranks (PR) of each of the three parts of the GRE (Biderman, 2013).

Our prediction model using Biderman's Formula Score has two predictors in addition to Biderman's Formula Score: undergraduate grade point average as a stand-alone variable and the student took calculus as an undergraduate. An astute observer might criticize this model for incorporating undergraduate grade point average twice, once as an individual factor and a second time as part of Biderman's Formula Score. The justification for its inclusion both times is for assessing multicollinearity among the variables. The statistics show that there was very little overlap of the predictors, signifying there is little adverse effect on the model.

As occurred in the previous model, the receiver operating characteristic analysis demonstrates that any combination of two or more of the predictors was the cut-point (Figure 5.4). Any student with a combination of any two of these three predictors or all three of the predictors has almost twenty-one times greater chance of being successful in a graduate athletic training program compared to a student who has either one or none of the predictors. Stated another way, the relative frequency of success found a student with any combination of two or all three of the predictors is almost more than twice as likely to be successful in the graduate athletic training program compared to a student who does not have one or none of the predictors.

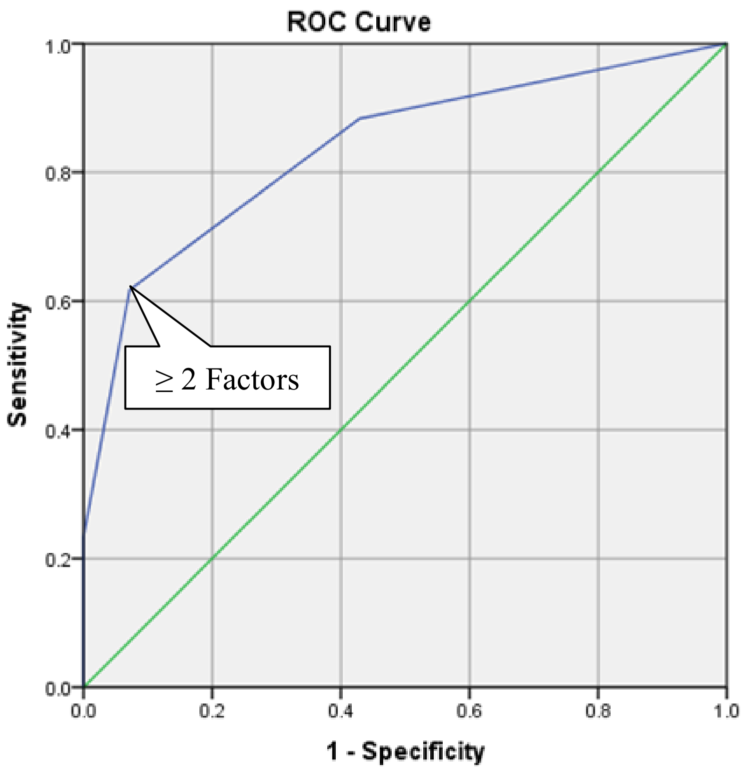

Figure 5.3 Receiver operating characteristic curve with identification of the optimum cut-point for the number of positive factors for prediction of success in the graduate athletic training program as indicated by graduate grade point average at the end of the first year ≥ 3.45 (includes Biderman's Formula Score)

Although the relationship of having any combination of two of the three predictors is very strong, it does not explain which combination of predictors is strongest. A series of analyses found students with both a high undergraduate grade point average and a high Biderman's Formula Score led to the greatest percentage of successful students in the graduate athletic training program (Figure 5.5). When adding the third predictor to the analysis, it was best to have a high Biderman's Formula Score, high undergraduate grade point average, and to have taken calculus. In this study, everybody who had all three of these criteria was successful all of the time. However, if the student's Biderman's Formula Score was low, but their undergraduate grade point average was high and they took calculus, they were still very successful too.

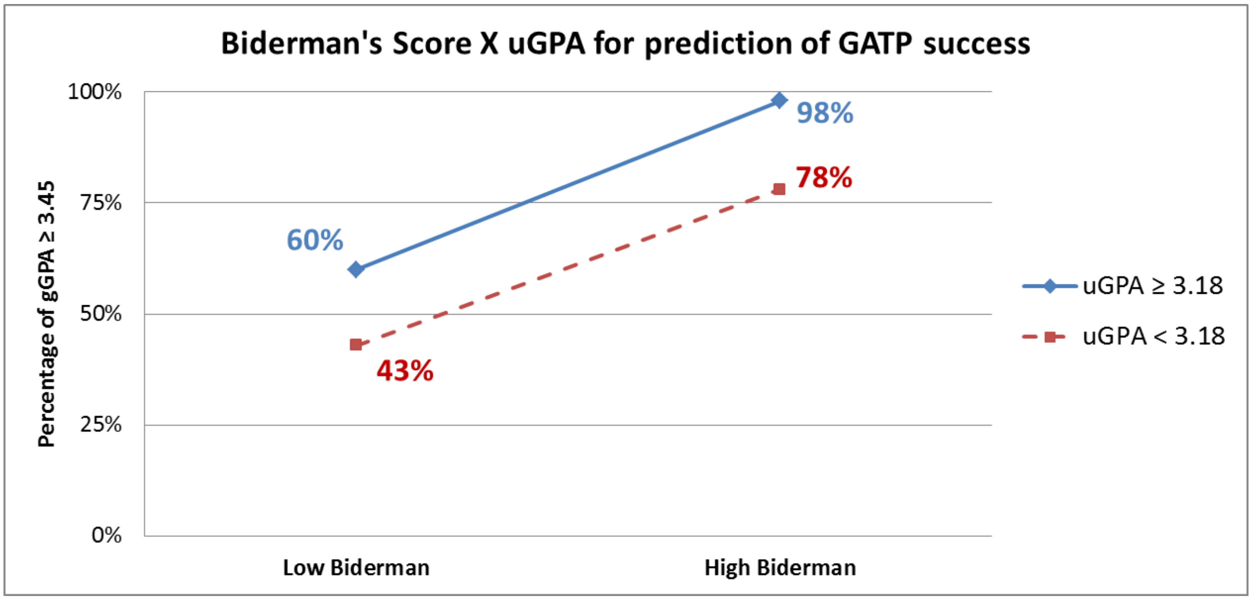

Figure 5.4 Possible interaction of Biderman's Formula Score X undergraduate grade point average (uGPA) for the prediction of graduate athletic training program success (graduate GPA at the end of the first year ≥ 3.45)

Prediction with Class of 2014 – Initial Prediction Model

The next class of eligible students to take the Board of Certification exam will be in the spring 2014 (after the completion of this study). An analysis of the students in the class of 2014 based on the initial three factor model (graduate grade point average at the end of the first-year, verbal score on the GRE, and the quantitative score on the GRE) and the number of predictor variables possessed by the students is shown in Table 5.1.

Table 5.1 Specific number of factors for a three factor model for prediction of first-attempt Pass versus Fail on the Board of Certification exam for the class of 2014 (GRE model)

Number of positive variables	Number of students with each number of variables
0	1
1	3
2	5
3	11
Total	20

Working Example for Predicting First-attempt Success on the Board of Certification Exam – Initial Model

As an example of how this model would work, we provide a set of students' data in Table 5.2. This table shows a series of students with the cut point for each of the predictors listed. If the student has a score at or above the cut-point it is listed in red. The far right column indicates the total number predictors the student possesses.

Table 5.2 Example of specific number of factors for the initial three-factor model for prediction of first-attempt Pass versus Fail on the Board of Certification exam

Student	**gGPA at the end of the first semester (≥ 3.45)**	**GREv (≥ 145.5)**	**GREq (≥ 143.5)**	**Total number of positive predictors**
Student #1	4.00	145	144	3
Student #2	3.40	143	145	1
Student #3	3.75	150	139	2
Student #4	4.00	146	150	3
Student #5	3.05	151	140	1
Student #6	3.00	140	142	0

Note. gGPA = Graduate Grade Point Average; GREv = Verbal section of the Graduate Record Examination; GREq = Quantitative section of the Graduate Record Examination;

Based on the data above, Students #1, #3, and #4 all have two or more of the three factors. According to the prediction model these three students have 6.3 times greater odds of passing the Board of Certification exam on their first-attempt compared to the odds of Students #2, #5, and #6 have of passing the Board of Certification exam on their first-attempt.

Prediction with Class of 2014 – Alternative Prediction Model

Using the same class data, (Class of 2014), an analysis of the students based on the alternative three-factor model (graduate grade point average at the end of the first-year and Biderman's Formula Score) and the number of predictor variables possessed by the students is shown in Table 5.3.

Table 5.3 Specific number of factors for a two-factor model for prediction of first-attempt Pass versus Fail on the Board of Certification exam for the class of 2014 (Biderman's Formula Score model)

Number of Positive Variables	Number of students with each number of variables
0	2
1	7
2	11
Total	20

Working Example for Predicting First-attempt Success on the Board of Certification Exam – Alternative Model

As an example of how this alternative model would work, we provide a set of students' data in Table 5.4. This table shows a series of students with the cut point for each of the predictors listed. If the student has a score at or above the cut-point it is listed in red. The far right column indicates the total number predictors the student possesses.

Table 5.4 Example of specific number of factors for the alternative prediction model for prediction of first-attempt Pass versus Fail on the Board of Certification exam

Student	gGPA at the end of the first semester (≥ 3.45)	[a]Biderman's Formula Score (≥ 420.5)	Total number of positive predictors
Student #1	4.00	465.0	2
Student #2	3.40	465.0	1
Student #3	3.75	444.0	2
Student #4	4.00	527.0	2
Student #5	3.05	397.0	0
Student #6	3.00	404.0	0

Note. gGPA = Graduate Grade Point Average

[a]Biderman's Formula Score = (100 * uGPA) + GREv PR + GREq PR + GREwr

Based on the data above, Students #1, #2, #3, and #4 all have at least one of the two predictors. According to the prediction model these students have 10.7 times greater odds of passing the Board of Certification exam on their first-attempt compared to the odds for Students #5 and #6 have of passing the Board of Certification exam on their first-attempt.

Comparison of the Models for Passing the Board of Certification Exam

In the first model, Students #1, #3, and #4 met the criteria for prediction of passing the Board of Certification exam. In the second model, these same students were predicted to be successful along with Student #2. The difference, which is not shown here is that this student had a very strong GRE analytical written score, (4.5 out of 6.0, which translates to a percentile rank of 72) (Educational Testing Services, 2011a). Although Student #2 was predicted to not be successful on the Board of Certification exam in the first model, this same student was predicted to be successful in the second model.

A logical question would be to ask which model should be followed. The answer depends on what one is looking for: an easier model to use or a model which gives a more complete picture of the individual's academic credentials, but requires a calculation to be performed. According to the outcome measures, the second prediction model (uGPA ≥ 3.18 and Biderman's Formula Score ≥ 420.5) produced an odds ratio of 10.7 and a relative frequency of success of 2.05. The results for the Class of 2014 remain to be seen; consequently, outside the scope of this specific study.

Examination of 2013 Recruiting Class

A total of 101 candidates expressed interest in the graduate athletic training program, but only 64 candidates had complete data sets to use in this analysis. From this group of 64 potential candidates, 23 candidates were offered positions to the graduate athletic training program. There were 16 students who accepted the offer to join the 2013 cohort, while two additional students were offered positions, but chose to defer their place in the program for one-year for personal reasons. Decisions on who to accept or not to accept into the graduate athletic training program were made prior to the prediction models found in this study were discovered. The comparison of those candidates offered a position in the graduate athletic training program (n = 23) to those candidates who were not offered a position (n = 41) in terms of predicting who would be successful in the graduate athletic training program based on the initial prediction model (undergraduate grade point average, Graduate Record Exam quantitative score, and the student took calculus as an undergraduate) is found in Table 5.5.

Table 5.5 Summary of positive factors (GRE Model) possessed by applicants to the graduate athletic training program (GATP) for the cohort 2013 to predict success in the graduate athletic training program

Number of Predictors	Offered a position in the GATP	Not offered a position in the GATP	Total	Percentage with number of predictors
0	0	4	4	6.25%
1	3	14	17	26.56%
2	11	19	30	46.88%
3	9	4	13	20.31%
Total	23	41	64	

Working Example for Predicting Success in the Graduate Athletic Training Program of Potential Candidates – Initial Model

As an example of how this model would work for the initial set of predictors for success in the graduate athletic training program, we provide a set of candidates' data in Table 5.6. This table shows a series of students with the cut point for each of the predictors listed. If the student has a score at or above the cut-point it is listed in red. The far right column indicates the total number predictors the student possesses.

Table 5.6 Example of specific number of factors for the initial three-factor model for predicting success in the graduate athletic training program, based on candidates' application data

Candidate	**uGPA ($\geq$ 3.18)**	**GREq ($\geq$ 141.5)**	**Student took calculus as an undergraduate (Yes or No)**	**Total number of positive predictors**
Student #7	3.20	144	Yes	3
Student #8	3.32	138	No	1
Student #9	3.11	142	No	1
Student #10	3.89	145	No	2
Student #11	3.43	156	No	2
Student #12	3.05	141	Yes	1
Student #13	3.68	147	No	2
Student #14	3.97	151	No	2
Student #15	4.00	156	Yes	3
Student #16	2.86	132	Yes	1

Note. uGPA = Undergraduate Grade Point Average; GREq = Quantitative section of the Graduate Record Examination

Based on the data above, Students #7, #10, #11, #13, #14, and #15 all have two or more of the predictors. According to the prediction model, these six students have almost 18 times greater odds of being successful in the graduate athletic training program than Students #8, #9, #12, and #16 have of being successful in the graduate athletic training program. Furthermore, if these data had been used for criteria for admission decisions on who is offered a position in the graduate athletic training program, offers would be made to the six candidates with two or more of the predictors.

Working Example for Predicting Success in the Graduate Athletic Training Program of Potential Candidates for Accepted to the Graduate Athletic Training Program – Alternative Model

As an example of how this model would work for the alternative set of predictors for success in the graduate athletic training program, we provide a set of candidates' data in Table 5.7.

Table 5.7 Example of specific number of factors for the alternative three-factor model for predicting success in the graduate athletic training program, based on candidates' application data

Candidate	**uGPA (≥ 3.18)**	**aBiderman's Formula Score (≥ 458.45)**	**Student took calculus as an undergraduate (Yes or No)**	**Total number of positive predictors**
Student #7	3.20	402.0	Yes	2
Student #8	3.32	394.0	No	1
Student #9	3.11	367.0	No	0
Student #10	3.89	445.0	No	1
Student #11	3.43	467.0	No	2
Student #12	3.05	470.0	Yes	2
Student #13	3.68	499.0	No	2
Student #14	3.97	485.0	No	2
Student #15	4.00	615.0	Yes	3
Student #16	2.86	438.0	Yes	1

Note. uGPA = Undergraduate Grade Point Average; .

[a]Biderman's Formula Score = (100 * uGPA) + GREv PR + GREq PR + GREwr

Based on the data above, Students #7, #11, #12, #13, #14, and #15 all have two or more of the predictors. According to the prediction model, these six candidates have almost 21 times greater odds of being successful in the graduate athletic training program than Students #8, #9, #10, and #16. Furthermore, if these data were used for criteria for admission decisions on who is offered a position in the graduate athletic training program, offers would be made to the six candidates with two or more of the predictors.

Comparison of the Models for Success in the Graduate Athletic Training Program

In the first model, Students #7, #10, #11, #13, #14, and #15 met the criteria for the prediction of success in the graduate athletic training program. In the second model, there was a slight change in which students would be predicted for success in the graduate athletic training program as Students #7, #11, #12, #13, #14, and #15 met the criteria. Student #10 met the prediction criteria for the initial model based on the strength of the undergraduate grade point average and their GRE quantitative score. But when the percentile rank scores from all three parts of the GRE are used for Biderman's Formula Score this student drops from the group of predicted to be successful. Student #12 was predicted to be successful in the alternative model based on a strong Biderman's Formula Score and he or she took calculus. Although this student's GRE quantitative score was just below the cut-point, for the initial model, their other GRE scores when used in Biderman's Formula Score were strong enough to provide this student with a second factor and place them in the group to be predicted successful in the graduate athletic training program. Experience has taught the selection committee when a student has a low undergraduate grade point average, but has a strong set of GRE scores to examine the student's entire body of work. Although it did not prove to be a reliable and valid measure across this sample, (likely due to a small sample size), students with this profile also tend to have taken more of the hard sciences and advanced math courses as an undergraduate, (i.e., calculus and physics, which physics was one of the final factors to drop out of the logistic regression).

Applicability to other Graduate Athletic Training Programs

The population from which the sample was used came from one specific graduate athletic training program. Although the use of graduate grade point average may be confirmed as a predictor of both success in other graduate athletic training programs, and of first-attempt Board of Certification exam success, it is not likely that all graduate athletic training programs will have a graduate grade point average cut-point equivalent to 3.45 as was determined and used for this study. In order for these results to have utility in the athletic training profession, two final prediction models were produced. All of the previously used predictor variables, except for graduate grade point average, were entered into another logistic regression analysis.

The results of this examination found two potential models using GRE scores and not Biderman's Formula Score: one a three-factor model including GRE verbal score, GRE quantitative score, and GRE analytical written score, and a second two-factor model with only GRE verbal score and GRE quantitative score. The three-factor model produced the strongest set of predictors for first-attempt success on the Board of Certification exam with any combination of two or more of the three variables yielding an odds ratio of 10.69 times and an relative frequency of success of 2.05. The regression analysis was repeated using Biderman's

Formula Score instead of the GRE scores directly, and the outcome yielded a model in which the only predictor was Biderman's Formula Score, and this model produced an odds ratio of 4.78.

Board of Certification exam and graduate athletic training program success

There are three different facets to athletic training education. The first is the education curriculum. Each athletic training education program is accredited by the Commission on Accreditation of Athletic Training Education (CAATE). Receiving accreditation certifies the athletic training education program is able to provide the requisite educational experience to prepare students to sit for the Board of Certification exam (Commission on Accreditation of Athletic Training Education, 2013a), which is the second part of athletic training education.

The certification exam is created and administered by the Board of Certification in order to test one's skills and knowledge as an entry-level athletic trainer (Board of Certification, 2011a; Ebel, 1999). "The purpose of the Board of Certification exam is to protect the public by ensuring that candidates for certification have achieved entry level competence" (Board of Certification, 2013a, p. 13). A 1978 article from *Athletic Training – The Journal of National Athletic Trainers' Association* outlined the first-attempt pass rate during the initial seven years of the administration of the certification exam. The authors cite a first-attempt pass rate of 91%, and go on to state:

> A number of failing candidates have been successfully reexamined and others have failed repeatedly to meet the high standards of the certification board. Those individuals should not be embarrassed by this failure since certification is recognition only of the highest level of competence in this field. (Westphalen & McLean, 1978, p. 91)

The third part of athletic training education is continuing education, which is also under the purview of the Board of Certification. Starting in 2014, a new standard for continuing education will be implemented whereby each certified athletic trainer must earn 50 hours every two years. Additionally, athletic trainers must maintain their certification in emergency cardiac care (Board of Certification, 2013b). The intent of continuing education is to “promote continued competence” in the knowledge and skill of an athletic trainer (Board of Certification, 2013b, "New Definition of CE" box).

The common characteristic among all three components of athletic training education is a desire to produce competent athletic trainers. Competence by definition is having basic skills or knowledge in some area or discipline (The Free Dictionary by Farlex, 2000a). In other words, the goal of CAATE accredited athletic training education programs and the Board of Certification is to produce and maintain professionals who have basic skills and knowledge in athletic training.

The term proficient means to have a level of understanding, knowledge or skill beyond competence (The Free Dictionary by Farlex, 2010). A search of both the Board of Certification and CAATE web sites for the words proficient or proficiency yielded no results. Many athletic training education programs focus solely on preparing and having their students pass the Board of Certification exam. With the new standard stating a school must have a pass rate of 70% or higher to be in compliance, (Commission on Accreditation of Athletic Training Education, 2013a) this focus will potentially increase.

Undergraduate education is intended to provide students with a wide breadth of experiences and education. There are few health professions that do not require graduate level education. The purpose of graduate education is to provide advanced or specialized curriculum in a discipline or profession. This in-depth education is intended to provide the student opportunities to become an expert in their chosen area of study. Additionally, graduate school provides students occasions to engage in higher-order learning and thinking, problem-solving, critical thinking, written and oral expression, and the utilization of technology as they applies to their particular profession (Pasco, 2009). Stated differently, the purpose of graduate school is to help create proficient clinicians and professionals. By identifying those candidates who possess qualities which are potential indicators of likely success in a graduate athletic training program, the goal for a program would be to seek clinicians who will not only be competent, (i.e., pass the Board of Certification exam on their first attempt), but will strive further to become proficient professionals.

Limitations and Future Research

The sample used for this research came from one specific graduate athletic training program. In some cases this led to small cell counts when the data were divided into various strata causing unstable results and large confidence intervals. To further validate the prediction models produced in this research, the next logical step is to apply them to other graduate athletic training programs or combine these data with like data from other graduate athletic training programs.

A final component of any prediction model is to conduct an impact analysis such as examining the economic effect the model has upon the associated population is indicated (Bruce & Wilkerson, 2010a; Childs & Cleland, 2006). Future studies examining the impact could be done. These studies could not only examine the financial impact upon students taking the Board of Certification exam multiple times, but studying the personal earning potential upon students predicted to be successful versus those predicted not to be successful in the graduate athletic training program in terms of the initial salaries or changes in their financial situations over a determined period of time.

Studies examining earning potentials have been conducted in the past. Generally speaking, there is already evidence that “individuals with a bachelor’s degree earn 50% more during their lifetime than . . . individuals with . . . (only) a high school diploma” (Barrow, Brock, & Rouse, 2013, p. 5). There is also evidence that individuals in the health support professions earn less than their STEM (Science, Technology, Engineering, and Mathematics) counterparts (Oreopoulos & Petronijevic, 2013). Although Oreopoulos & Petronijevic (2013) did not define specifically what qualified as a “health support profession,” athletic training can be classified in such a category. In a 2010 study, they found “college graduates in the health professions earned about 68% more on average than high school graduates in the health professional sector” (Oreopoulos & Petronijevic, 2013, p. 46). However, those “college graduates” in the health support professions earned only 27% more than those with only a high school diploma in the health support professions (Oreopoulos & Petronijevic, 2013, pp. 45-46). A 2011 salary survey conducted by the National Athletic Trainers’ Association found those athletic trainers with a Master’s degree earned about $5000 more per year than athletic trainers with only a Bachelor’s degree (Lowe, 2011). What has not been studied specifically is the starting salary of graduates from a graduate athletic training program versus graduates of an undergraduate athletic training program, since they both would enter the profession with no experience as a certified athletic trainer.

Identifying students who are both likely to be successful in a graduate athletic training program, and who are likely to pass the Board of Certification exam on their first attempt, may indirectly identify students who are likely to remain in the athletic training profession versus pursuing other allied health professions. Such identification may make it less likely students will get toward the end of the educational process only to decide that athletic training is not for them. These students have invested considerable time, energy, and money in their education only to find they are "stuck" with few options. This predicament results in a waste of considerable resources for all involved. Research to investigate if early identification of potentially successful students results in a long term commitment to the athletic training profession would be valuable.

Applying the methods of creating prediction models to other allied health professions such as physical therapy, occupational therapy, nursing, etc., would yield potentially interesting data and results. None of the procedures, methods, or information used to generate these prediction models is exclusive to athletic training. All of the information available can be gathered through standard data collection methods from graduate school application files. Variables and cut-points might differ across professions, but how those associated data and predictors are generated would remain consistent.

Biderman's Formula Score had only been utilized in the University of Tennessee at Chattanooga Psychology Department. Previous research utilizing the methods described in this study has not be conducted (Biderman, 2013). Specific studies to examine its reliability and validity across other programs and institutions should also be investigated.

Previous attempts to predict first-attempt Board of Certification exam success were not successful for a variety of reasons. Each of the previous studies used frequentist statistics where this dissertation utilized Bayesian philosophy. A study examining potential predictors for first-attempt Board of Certification exam success at the undergraduate level utilizing the methods implemented in our study may produce successful prediction models at the undergraduate level of athletic training education.

A limitation discussed in Chapter I was the effort given by students on the GRE because the graduate athletic training program did not have minimum GRE score requirements. With the data generated from these prediction models, and communication of their results, it is reasonable to expect potential students to take the GRE more seriously; thus a potential increase in the scores may be a result. There is a likelihood the calculations outlined here may need to be revised periodically to reflect an increase in the quality of the students applying to the graduate athletic training program.

Clinical Relevance

The Commission on Accreditation of Athletic Training Education accreditation standards require all athletic training education programs to demonstrate a three-year aggregate first-time pass rate of 70% (Commission on Accreditation of Athletic Training Education, 2013a). Programs will be forced to place greater emphasis of passing the Board of Certification exam on the first attempt. Consequently, programs will need to be able to identify students who are most likely to pass the Board of Certification exam on the first attempt. This study has provided a blueprint for accomplishment of this task.

The significance of these results is timely. The Executive Committee for Education (ECE) of the NATA is in the process of exploring the most appropriate professional degree for athletic trainers to be eligible to sit for the Board of Certification exam. In 2012, the ECE published a white paper entitled the *Future Direction in Athletic Training Education* (Brown, 2012) which includes 14 different recommendations. The second recommendation has created the strongest passions and debate among the membership:

> Recommendation #2: The NATA, with support from the Strategic Alliance, should conduct a detailed analysis specifically focused on professional education in athletic training that will be completed by June 2014. A key outcome of this analysis will be a determination of the most appropriate professional degree to position athletic trainers to provide positive patient outcomes and ensure the longevity of the profession of athletic training. (Brown, 2012, p. 2)

Presently, the debate throughout the athletic training profession is whether or not a Master's degree should be the minimum requirement in order for a student to sit for the Board of Certification exam. There are approximately 350 accredited athletic training education programs, of which 26 are graduate professional (entry-level) athletic training education programs (Commission on Accreditation of Athletic Training Education, 2013d). Several athletic training education curricula are in the process of converting from the undergraduate model to the graduate professional (entry-level) athletic training education program in anticipation of the direction professional education appears to be moving (Commission on Accreditation of Athletic Training Education, 2013c).

Although the final recommendations from the ECE have not been made to the NATA Board of Directors, much discussion has taken place regarding the direction the profession should take for a minimum academic degree to be eligible to sit for the Board of Certification exam. Many who have expressed their concern over moving to a graduate professional (entry-level) athletic training education program have a background rooted in the undergraduate curricula, so an obvious bias appears to exist in their writings (Grantham, 2013; Hauth, 2012; Henning, 2012; Hooker, 2013; Meyer, 2013; Pitney, 2012; Prentice, 2013). Only one article has been published in favor of the graduate professional (entry-level) athletic training education program from faculty who have experience in both undergraduate and graduate education programs (Wilkerson, Colston, & Bogdanowicz, 2006). A strong argument was made that graduate professional (entry-level) athletic training education is needed to advance the profession. The significant role the GRE has in all of the prediction models producing large odds ratios, and significant relative frequency of success values, cannot be discounted. None of the previous studies attempting to predict Board of Certification success at the undergraduate level were successful, and to be eligible to take the GRE a student must be near the end of baccalaureate studies. Hence, converting to graduate professional (entry-level) athletic training education programs for Board of Certification eligibility makes the most sense.

The single point all individuals seem to agree upon is the clear need for substantial change, but pursuing a graduate professional (entry-level) athletic training education program as the only route to certification has many concerned and fearful about what might happen after implementation of such a requirement. Should a mandated conversion from undergraduate athletic training education to a graduate professional (entry-level) athletic training education program be issued, then the results of this study will likely be valued by program directors. A likely future goal of athletic training program directors will be to identify objective methods to use in their search to identify those students who are likely to pass the Board of Certification exam on their first attempt, are likely to be successful in their graduate athletic training programs and become proficient professionals after graduation.

Conclusion

The prediction models created for identifying students likely to pass the Board of Certification exam on their first attempt and for identifying students who will be successful in a graduate athletic training program generated very strong odds ratios. The predictors associated with success were related to past academic performance either through grade point average, GRE performance, or that the student took calculus as an undergraduate. A very strong predictor which incorporates both undergraduate grade point average and GRE (PR) scores was Biderman's Formula Score. With the increased demands by the accrediting body for a minimum Board of Certification exam pass rate to be in compliance, and with a potential shift to a graduate professional (entry-level) athletic training education as the entry-point to sit for the Board of Certification exam, the methods for the generation of the specific prediction models created in this study will have potential uses throughout not only athletic training, but other professions too.

REFERENCES

About.com. (2013). About.com College Admissions Retrieved August 26, 2013, from http://collegeapps.about.com/

Ardern, C. L., Taylor, N. F., Feller, J. A., Whitehead, T. S., & Webster, K. E. (2013). Psychological responses matter in returning to preinjury level of sport after anterior cruciate ligament reconstruction surgery. *American Journal of Sports Medicine, 41*(7), 1549-1558. doi: 10.1177/0363546513489284

Armstrong, A., Dahl, C., & Haffner, W. (1998). Predictors of performance on the National Board of Medical Examiners obstetrics and gynecology subject examination. *Obstetrics and Gynecology, 91*(6), 1021-1022.

Bailey, R. T. (2002, April 4-5). *Shouldn't I get an "A?".* Paper presented at the 2002 ASEE Southeastern Section Annual Meeting, Gainesville, FL.

Baldwin, B., & Bruce, S. L. (2008). *Core stabilization and shoulder dysfunction in collegiate softball players*. Paper presented at the Orthopedic Research Grand Rounds, Chattanooga, TN.

Balogun, J. A., Karacoloff, L. A., & Farina, N. T. (1986). Predictors of academic achievement in physical therapy. *Physical Therapy, 66*(6), 976-980.

Barrow, L., Brock, T., & Rouse, C. E. (2013). Postsecondary education in the United States: Introducing the issue. *The Future of Children, 23*(1), 3-16.

Beneciuk, J. M., Bishop, M. D., & George, S. Z. (2009). Clinical prediction rules for physical therapy interventions: A systematic review. *Physical Therapy, 89* (2), 114-124.

Biderman, M. D. (2013). Admission criteria: The formula score Retrieved January 12, 2013, from http://www.utc.edu/Academic/Industrial-OrganizationalPsychology/AdmissionCriteria.php

Billings, C. E. (2004). Epidemiology of injuries and illnesses during the United States Air Force Academy 2002 basic cadet training program: documenting the need for prevention. *Military Medicine, 169*(8), 664-670.

Board of Certification (BOC) Certification Examination for Athletic Trainers. (2008). 2007 Annual summary. Princeton, NJ.

Board of Certification (BOC) certification examination for athletic trainers. (2009). Annual summary for 2008 testing year. Princeton, NJ.

Board of Certification, I. (2011a). Exam development & scoring Retrieved July 29, 2012, from http://www.bocatc.org/educators/exam-development-scoring

Board of Certification, I. (2011b). Role delineation study/practice analysis: Blueprint for the exam and recertfication activities (6th ed.). Omaha, NE: National Athletic Trainers' Association Board of Certification.

Board of Certification, I. (2013a). BOC exam candidate handbook: A step by step guide for candidates preparing for the BOC exam Retrieved October 15, 2013, from http://bocatc.org/images/stories/candidates/boc_candidate_handbook_1305cf.pdf

Board of Certification, I. (2013b). Maintain certification - 2014-2015 Changes Retrieved October 15, 2013, from http://bocatc.org/ats/maintain-certification

Böhning, D., Böhning, W., & Holling, H. (2008). Revisiting Youden's index as a useful measure of the misclassification error in meta-analysis of diagnostic studies. *Statistical Methods in Medical Research, 17*(6), 543-554.

Brenner, A. K. (2008). Clinical prediction rule for those soldiers most likely to develop lower extremity stress fractures during initial entry training. In: Program and abstracts of the 2008 American Physical Therapy Association combined sections meeting. Nashville, Tennessee, February 6-9, 2008. *Journal of Orthopaedic & Sports Physical Therapy, 38*(1), A76.

Bretz, R. D., Jr. (1989). College grade point average as a predictor of adult success: a meta-analytic review and some additional evidence. *Public Personnel Management*, *18*(1), 11.

Brown, S. (2012). Future directions in athletic training education. Dallas, TX: NATA Executive Committee for Education.

Bruce, S. L. (2011). *[Passing rates on the BOC exam, UTC vs. national passing rates]*.

Bruce, S. L. (2012). *How to create useful clinical prediction rules: Predicting who is likely to get injured & the prevention domaiin*. Paper presented at the 63rd National Athletic Trainers' Association Annual Meeting and Clinical Symposium, St. Louis, MO. http://members.nata.org/annualmeeting/HandoutLibrary/2012/HandoutLibrary.cfm?docToServe=Assessing-the-Prevention-Domain.pdf

Bruce, S. L., & Wilkerson, G. B. (2010a). Clinical prediction rules, part 1: Conceptual overview. *Athletic Therapy Today, 15*(2), 4-9.

Bruce, S. L., & Wilkerson, G. B. (2010b). Clinical prediction rules, part 2: Data analysis procedures and clinical application of results. *Athletic Therapy Today, 15*(2), 10-13.

Burdette, R. N., & Wilkerson, G. B. (2012). *Pre-season characteristics as predictors of musculoskeletal injury risk*. Paper presented at the Graduate Research Day, University of Tennessee at Chattanooga, Chattanooga, TN.

Burton, N. W., & Wang, M.-m. (2005). Predicting long-term success in graduate school: A collaborative validity study (pp. 70). Princeton, NJ Educational Testing Service

Casella, G. (2008). *Refresher on Bayesian and frequentist concepts: Models, assumptions, and inference*. Paper presented at the American College Clinical Pharmacology, Philadelphia, PA.

CASTLE Worldwide, I. (2001). Annual report for the 2000 testing year. Princeton, NJ.

Childs, J. D., & Cleland, J. A. (2006). Development and application of clinical prediction rules to improve decision making in physical therapist practice. *Physical Therapy, 86*(1), 122.

Childs, J. D., Fritz, J. M., Flynn, T. W., Irrgang, J. J., Johnson, K. K., Majkowski, G. R., & Delitto, A. (2004). A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: A validation study. *Annals of Internal Medicine, 141*(12), 920-928.

Childs, J. D., Fritz, J. M., Piva, S. R., & Erhard, R. E. (2003). Clinical decision making in the identification of patients likely to benefit from spinal manipulation: A traditional versus an evidence-based approach. *Journal of Orthopaedic & Sports Physical Therapy, 33*(5), 259-272.

Clark, A. S., Bruce, S. L., & Wilkerson, G. B. (2012). *The relationship between neurocognitive reaction time and incidence of core or lower extremity sprains or strains*. Paper presented at the Orthopedic Research Grand Rounds, Chattanooga, TN.

Clark, E. L. (1964). Reliability of grade point averages. *The Journal of Educational Research, 57*(8), 428-430.

Cleland, J. A., Childs, J. D., Fritz, J. M., Whitman, J. M., & Eberbart, S. L. (2006). A clinical prediction rule for classifying patients with neck pain who demonstrate short-term improvement with thoracic spine thrust manipulation [abstract]. *Journal of Manual Manipulative Therapy, 14*(3), 171-172.

Cleland, J. A., Childs, J. D., Fritz, J. M., Whitman, J. M., & Eberhart, S. L. (2007). Development of a clinical prediction rule for guiding treatment of a subgroup of patients with neck pain: use of thoracic spine manipulation, exercise, and patient education. *Physical Therapy, 87*(1), 9-23.

Cockrell, K. N., & Bruce, S. L. (2008). *Predictors of lower extremity injury in female collegiate soccer and softball athletes*. Paper presented at the Orthopedic Research Grand Rounds, Chattanooga, TN.

Cohen-Schotanus, J., Muijtjens, A. M. M., Reinders, J. J., Agsteribbe, J., van Rossum, H. J. M., & van der Vleuten, C. P. M. (2006). The predictive validity of grade point average scores in a partial lottery medical school admission system. *Medical Education, 40*, 1012-1019. doi: 10.1111/j.1365-2929.2006.02561.x

Commission on Accreditation of Athletic Training Education. (2013a). Commission on Accreditation of Athletic Training Education Retrieved October 14, 2013, from http://www.caate.net/

Commission on Accreditation of Athletic Training Education. (2013b, August 20, 2013). Posting of outcomes. *Becoming an athletic trainer* Retrieved September 8, 2013, from http://caate.occutrain.net/?s=pass+rate

Commission on Accreditation of Athletic Training Education. (2013c). Professional program updates. Round Rock, TX: CAATE.

Commission on Accreditation of Athletic Training Education. (2013d). Search for accredited programs Retrieved September 15, 2013

Common Data Set Initiative. (2012). Common daat set 2012-2013 Retrieved July 14, 2013, from http://www.commondataset.org/

Concato, J., Feinstein, A. R., & Holford, T. R. (1993). The risk of determining risk with multivariable models. *Annals of Internal Medicine, 118*(3), 201-210.

Covey, S. R. (2004). *The 7 habits of highly effective people: Restoring the character ethis*. New York, NY: Free Press.

Craig, D. I. (2003). Educational reform in athletic training: a policy analysis. *Journal of Athletic Training, 38*(4), 351.

Daehnert, C., & Carter, J. D. (1987). The prediction of success in a clinical psychology graduate program. *Educational and Psychological Measurement, 47*(4), 1113-1125. doi: 10.1177/0013164487474029

Darling-Hammond, L., & Rustique-Forrester, E. (2005). The consequences of student testing for teaching and teacher quality. *Yearbook of the National Society for the Study of Education, 104*(2), 289-319.

Davenport, T. E., Cleland, J., & Kulig, K. (2009). Patient classification based on psychosocial variables predicts treatment outcomes in patients with lower back pain who meet a clinical prediction rule. *Journal of Orthopaedic & Sports Physical Therapy, 39*(1), A19-A20. Abstract OPL17.

Day, J. A. (1986). Graduate Record Examination analytical scores as predictors of academic success in four entry-level master's degree physical therapy programs. *Physical Therapy, 66*(10), 1555.

de Virgilio, C., Yaghoubian, A., Kaji, A., Collins, J. C., Deveney, K., Dolich, M., . . . Liu, T. (2010). Predicting performance on the American Board of Surgery qualifying and certifying examinations: a multi-institutional study. *Archives of Surgery*, *145*(9), 852-856.

DeAngelis, S. (2003). Noncognitive predictors of academic performance: Going beyond traditional measures. *Journal of Allied Health, 32*, 52-57.

Delforge, G. D., & Behnke, R. S. (1999). The history and evolution of athletic training education in the United States. *Journal of Athletic Training*, *34*(1), 53-61.

Denegar, C. R. (2012). *How to objectively assess the prevention domain: Applying clinical prediction rules (guides?) uses in healthcare*. Paper presented at the 63rd National Athletic Trainers' Association Annual Meeting and Clinical Symposium, St. Louis, MO. http://members.nata.org/annualmeeting/HandoutLibrary/2012/HandoutLibrary.cfm?docToServe=Clinical-Prediction-Rules.pdf

Denegar, C. R., & Cordova, M. L. (2012). Application of Statistics in Establishing Diagnostic Certainty. *Journal of Athletic Training, 47*(2), 233.

Denegar, C. R., & Wilkerson, G. W. (2013). *Bridging the chasm between research & clinical practice in athletic training: A discussion of methods and analysis*. Paper presented at the 2013 National Athletic Trainers' Association Educator's Conference, Dallas, TX.

Draper, D. O. (1989). Students' learning styles compared with their performance on the NATA certification exam. *Athletic Training - The Journal of the National Athletic Trainers' Association*, *24*(3), 234-235, 275.

Ebel, R. G. (1999). *Far beyond the shoe box: Fifty years of the National Athletic Trainers' Association*. New York, NY: Forbes Custom Publishing.

Educational Testing Services. (2011a). GRE 2011-2012: Guide to the use of scores Retrieved June 17, 2012, from http://www.ets.org/s/gre/pdf/gre_guide.pdf

Educational Testing Services. (2011b). Purpose of standardized tests Retrieved June 17, 2012, from http://www.ets.org/understanding_testing/purpose/

Educational Testing Services. (2013a). Analytical writing interpretive data used on score reports Retrieved May 21, 2012, from http://www.ets.org/s/gre/pdf/gre_guide_table1a.pdf

Educational Testing Services. (2013b). Verbal reasoning and quantitative reasoning interpretative data used on score reports Retrieved May 21, 2012, from http://www.ets.org/s/gre/pdf/gre_guide_table1a.pdf

Emparanza, J. I., & Aginaga, J. R. (2001). Validation of the Ottawa Knee Rules. *Annals of Emergency Medicine*, *38*(4), 364-368.

Erickson, M. A., & Martin, M. (2000). Contributors to initial success on the National Athletic Trainers' Association Board of Certification Examination as perceived by candidate sponsors: A delphi study. *Journal of Athletic Training, 35*(2), 134-138.

Etaugh, A. F., Etaugh, C. F., & Hurd, D. E. (1972). Reliability of college grades and grade point averages: Some implications for prediction of academic performance. *Educational and Psychological Measurement, 32*, 1045-1050.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27* 861-874.

Federation of State Boards of Physical Therapy. (2012). Welcome to the FSBPT website Retrieved June 10, 2012, from https://www.fsbpt.org/index.asp

Feinstein, A. R. (1996). *Multivariable analysis: An introduction*. New Haven, CT: Yale University Press.

Feldman, A. (2007). *Are GPA and standardized test scores significant predictors of success for occupational therapy students.* (M.S.O.T.), Touro University.

Ferguson, E., James, D., & Madeley, L. (2002). Factors associated with success in medical school: systematic review of the literature. *British Medical Journal, 324*(7343), 952-957. doi: 910.1136/bmj.1324.7343.1952.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Thousasnd Oaks, CA: SAGE Publications.

Fienberg, S. E. (2006). When did Bayesian inference become "Bayesian"? *Bayesian Analysis, 1*(1), 1-40.

Flynn, T., Fritz, J., Whitman, J., Wainner, R., Magel, J., Rendeiro, D., . . . Allison, S. (2002). A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. *Spine*, *27*(24), 2835.

Friess, L., & Bruce, S. L. (2010). *Predictors of overuse shoulder injuries in collegiate volleyball athletes*. Paper presented at the Southeastern Athletic Trainers' Association 2nd Biennial Athletic Training Educators' Conference, Atlanta, GA.

George, S., Haque, M. S., & Oyebode, F. (2006). Standard setting: Comparison of two methods. *BMC Medical Education, 6*(1), 46-51.

Grace, P. (1999). Milestones in athletic trainer certification. *Journal of Athletic Training, 34*(3), 285-291.

Grantham, J. (2013). Athletic training education: What's next? *NATA News, 25,* 12-13.

Grossbach, A., & Kuncel, N. R. (2011). The predictive validity of nursing admission measures for performance on the National Council Licensure Examination: A meta-analysis. *Journal of Professional Nursing, 27,* 124-128.

Hamdy, H., Prasad, K., Anderson, M. B., Scherpbier, A., Williams, R., Zwierstra, R., & Cuddihy, H. (2006). BEME systematic review: Predictive values of measurements obtained in medical schools and future performance in medical practice. *Medical Teacher, 28*(2), 103-116.

Hansen, M. J., & Pozehl, B. J. (1995). The effectiveness of admission criteria in predicting achievement in a Master's degree program in nursing. *Journal of Nursing Education, 34*(9), 433-437.

The Americans with Disabilities Act of 1990, Title 42, chapter 126, of the United States Code beginning at section 12101 C.F.R. § 12101 et seq. (1990).

Harrelson, G. L., Gallaspy, J. B., Knight, H. V., & Leaver-Dunn, D. (1997). Predictors of success on the NATABOC certification examination. *Journal of Athletic Training, 32*(4), 323–327.

Haswell, K., Gilmour, J., & Moore, B. (2008). Clinical decision rules for identification of low back pain patients with neurologic involvement in primary care. *Spine, 33*(1), 68.

Hauth, J. M. (2012). Requiring professional athletic training programs at the post-baccalaureate level: Considerations and concerns - Athletic training education reform: Where is the cheese? *Athletic Training Education Journal, 7*(1), 7-10. doi: 10.5608/070104

Hayes, S. H., Fiebert, I. M., Carroll, S., R., & Magill, R. N. (1997). Predictors of academic success in a physical therapy program: Is there a difference between traditional and nontraditional students? *Journal of Physical Therapy Education, 11*(1), 10-16.

Haynes, R. B., Devereaux, P. J., & Guyatt, G. H. (2002). Clinical expertise in the era of evidence-based medicine and patient choice. *Evidence-Based Medicine, 7*(2), 36-38.

Henderson, J. P. (1998). Annual report on the National Certification Examination April 1997 through February 1998. Princeton, NJ.

Henley, S., Bruce, S. L., & McDermott, B. P. (2012). *Relationship of life events to injury risk in Division I-FCS college football players*. Paper presented at the Southeastern Athletic Trainers' Association 3rd Biennial Athletic Training Educators' Conference, Atlanta, GA.

Henning, J. (2012). Requiring professional athletic training programs at the post-baccalaureate level: Considerations and concerns - Invited Commentary. *Athletic Training Education Journal, 7*(1), 6-7. doi: 10.5608/070104

Hess, J. E., Wilkerson, G. B., & Colston, M. A. (2011). *Prediction of core muscle strains in NCAA Division I-FCS football players*. Paper presented at the Graduate Research Day University of Tennessee at Chattanooga, Chattanooga, TN.

Hess, J. E., Wilkerson, G. B., & Colston, M. A. (2012). *Prediction of core muscle strains in NCAA Division I-FCS football players*. Paper presented at the Southeastern Athletic Trainers' Association 3rd Biennial Athletic Training Educators' Conference, , Atlanta, GA.

Heyworth, J. (2003). Ottawa ankle rules for the injured ankle: Useful clinical rules save on radiographs and need to be used widely. *British Medical Journal, 326*(7386), 405.

Hickman, K. M. (2010). *Board of Certification examination success and clinical education.* (Doctor of Philosophy Dissertation), Virginia Polytechnic Institute and State University, Blacksburg, VA.

Hicks, G. E., Fritz, J. M., Delitto, A., & McGill, S. M. (2005). Preliminary development of a clinical prediction rule for determining which patients with low back pain will respond to a stabilization exercise program. *Archives of Physical Medicine and Rehabilitation, 86*(9), 1753-1762.

Hocking, J. A., & Piepenbrock, K. (2010). Predictive ability of the graduate record examination and its usage across physician assistant programs. *Journal of Physician Assistant Education, 21*(4), 18-22.

Hooker, D. N. (2013, June). Letter to the Editor. *NATA News, 25,* 8-9.

Horn, L. J., & Carroll, C. D. (1996). Postsecondary education descriptive analysis reports nontraditional undergraduates: Trends in enrollment from 1986 to 1992 and persistence and attainment among 1989-90 beginning postsecondary students: U.S. Department of Education, Office of Educational Research and Improvement.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Hubbard, D. W. (2010). *How to measure anything: Finding the value of intangibles in business*. Hoboken, NJ: John Wiley & Sons, Inc.

Hunsecker, J. G. (2007). *High stakes testing in Florida: Media portrayals and parental realities.* (Master of Arts), University of South Florida, Tampa, FL.

IBM Corporation. (2011). IBM SPSS Data Collection (Version 19.0) [Computer software]. Somers, NY: IBM.

Iverson, C. A., Sutlive, T. G., Crowell, M. S., Morrell, R. L., Perkins, M. W., Garber, M. B., . . . Wainner, R. S. (2008). Lumbopelvic manipulation for the treatment of patients with patellofemoral pain syndrome: Development of a clinical prediction rule. *Journal of Orthopaedic & Sports Physical Therapy, 38*(6), 297-312.

Johnson, S. B. (2010). Examination review for 2009-10 testing year: Board of Certification (BOC) certification examination for athletic trainers. Princeton, NJ.

Johnson, S. B. (2011). Examination review for 2010-2011 testing year: Board of Certification (BOC) certification examination for athletic trainers. Princeton, NJ.

Johnson, S. B. (2012). Examination report for 2011-2012 testing year: Board of Certification (BOC) certification examination for athletic trainers. Princeton, NJ.

Johnson, S. B. (2013). Examination report for 2012-2013 testing year: Board of Certification (BOC) certification examination for athletic trainers. Princeton, NJ.

Jones, M. M., Wilkerson, G. B., Colston, M. A., & Bruce, S. L. (2012). *Responses to the "Life Events Survey for Collegiate Athletes" as injury predictors*. Paper presented at the Orthopedic Research Grand Rounds, , Chattanooga, TN.

Julian, E. R. (2005). Validity of the Medical College Admission Test for predicting medical school performance. *Academic Medicine, 80* (10), 910-917.

Karch, E., Wilkerson, G. W., & Bruce, S. L. (2012a). *Analysis of college wrestlers' characteristics relevant to injury risk*. Paper presented at the Graduate Research Day University of Tennessee at Chattanooga, Chattanooga, TN.

Karch, E., Wilkerson, G. W., & Bruce, S. L. (2012b). *Analysis of college wrestlers' characteristics relevant to injury risk*. Paper presented at the Southeastern Athletic Trainers' Association 3rd Biennial Athletic Training Educators' Conference, Atlanta, GA.

Katz, J. R., Chow, C., Motzer, S. A., & Woods, S. L. (2009). The Graduate Record Examination: Help or hindrance in nursing graduate school admissions? *Journal of Professional Nursing, 25*(6), 369.

Keskula, D. R., Sammarone, P. G., & Perrin, D. H. (1995). Prediction of academic achievement in an NATA-approved graduate athletic training education program. *Journal of Athletic Training, 30*(1), 55-56.

Kirchner, G. L., & Holm, M. B. (1997). Prediction of academic and clinical performance of occupational therapy students in an entry-level master's program. *American Journal of Occupational Therapy, 51*(9), 775.

Kirchner, G. L., Holm, M. B., Ekes, A. M., & Williams, R. W. (1994). Predictors of student success in an entry-level master in physical therapy program. *Journal of Physical Therapy Education, 8*(2), 76.

Kosmahl, E. M. (2005). Factors related to physical therapist license examination scores. *Journal of Physical Therapy Education, 19*(2), 52.

Kreiter, C. D., & Kreiter, Y. (2007). A validity generalization perspective on the ability of undergraduate GPA and the Medical College Admission Test to predict important outcomes. *Teaching & Learning in Medicine, 19*(2), 95-100.

Kuijpers, T., van der Heijden, G. J. M. G., Vergouwe, Y., Twisk, J. W. R., Boeke, A. J. P., Bouter, L. M., & van der Windt, D. l., A. W. M. . (2007). Good generalizability of a prediction rule for prediction of persistent shoulder pain in the short term. *Journal of Clinical Epidemiology, 60*(9), 947-953.

Kuijpers, T., van der Windt, D. A. W. M., Boeke, A. J. P., Twisk, J. W. R., Vergouwe, Y., Bouter, L. M., & van der Heijden, G. J. M. G. (2006). Clinical prediction rules for the prognosis of shoulder pain in general practice. *Pain, 120* 276-285.

Kuncel, N. R., Crede', M., & Thomas, L. L. (2007). A meta-analysis of the predictive validity of the Graduate Management Admission Test (GMAT) and undergraduate grade point average (UGPA) for graduate student academic performance. *Academy of Management Learning & Education, 6*(1), 51-68.

Kuncel, N. R., & Hezlett, S. A. (2007). Standardized tests predict graduate students' success. *Science, 315* 1080-1081.

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the Graduate Record Examinations: Implications for graduate student selection and performance. *Psychological Bulletin, 127*(1), 162-181. doi: 10.1037/0033-2909.127.1.162

Kuncel, N. R., Wee, S., Serafin, L., & Hezlett, S. A. (2010). The validity of the Graduate Record Examination for master's and doctoral programs: A meta-analytic investigation. *Educational and Psychological Measurement, 70*(2), 340-353. . doi: 10.1177/0013164409344508

Lai, G. P., Mink, D. R., & Pasta, D. J. (n.d.). Beyond Breslow-Day: Homogeneity Across R x C Tables *Paper 74949* (pp. 1-7). San Francisco, CA: ICON Late Phase & Outcomes Research.

Laslett, M. (2006). Clinical prediction rule for rapid pain relief of low back pain following manipulation. *New Zealand Journal of Physiotherapy, 34*(2), 93.

Laupacis, A., Sekar, N., & Stiell, I. G. (1997). Clinical prediction rules. A review and suggested modifications of methodological standards. *Journal of the American Medical Association, 277*(6), 488-494. doi: 10.1001/jama.277.6.488

Leisey, J. (2004). Prospective validation of the Ottawa Ankle Rules in a deployed military population. *Military Medicine, 169*(10), 804.

Lesher, J. D., Sutlive, T. G., Miller, G. A., Chine, N. J., Garber, M. B., & Wainner, R. S. (2006). Development of a clinical prediction rule for classifying patients with patellofemoral pain syndrome who respond to patellar taping. *Journal of Orthopaedic & Sports Physical Therapy, 36*(11), 854-866.

Levine, S. B., Knecht, H. G., & Eisen, R. G. (1986). Selection of physical therapy students: Interview methods and academic predictors. *Journal of Allied Health, 15*(2), 143-151.

Lindquist, M., Arrington, S., & Scheopner, K. (2007). *The BOC Exam: The first 40 years: A tribute to our volunteers*. Lincoln, NE: Jacob North Printing Company.

Lowe, R. (2011). Athletictraining salaries on the rise according to latest survey. *NATA News, 23,* 12, 14.

MacDermid, J., & Law, M. (2007). Evaluating the evidence. In M. C. Law & J. MacDermid (Eds.), *Evidence-based rehabilitation: A guide to practice* (pp. 121-142). Thorofare, NJ: Slack Incorporated.

Mahieu, N. N., Witvrouw, E., Stevens, V., Van Tiggelen, D., & Roget, P. (2006). Intrinsic risk factors for the development of Achilles tendon overuse injury: A prospective study. *American Journal of Sports Medicine*, *34*(2), 226.

Masters, J. R. (1974). The relationship between number of response categories and reliability of Likert-type questionnaires. *Journal of Educational Measurement, 11*(1), 49-53.

McClintock, J. C., & Gravlee, G. P. (2010). Predicting success on the certification examinations of the American Board of Anesthesiology. *Anesthesiology, 112*(1), 212-219. doi: 210.1097/ALN.1090b1013e3181c1062e1092f

McGinnis, M. E. (1984). Admission predictors for pre-physical therapy majors. *Physical Therapy, 64*(1), 55.

McLean Jr., J. L. (1969). Does the National Athletic Trainers' Association need a certification examination? *Journal of the National Athletic Trainers' Association, 4*(1), 10-11.

Meleca, C. B. (1995). Traditional predictors of academic performance in a medical school's independent study program. *Academic Medicine, 70*(1), 59-63.

Melendez, A., Bruce, S. L., & Wilkerson, G. W. (2010). *Relationship of core fatigue-resistance to throwing accuracy as predictors factors for shoulder injuries in baseball players*. Paper presented at the Celebration of Graduate Student Scholarship, University of Tennessee at Chattanooga, Chattanooga, TN.

Mertler, C. A., & Vannetta, R. A. (2005a). *Advanced and Multivariate Statistical Methods* (3rd ed.). Los Angeles, CA: Pyrczak Publishing.

Mertler, C. A., & Vannetta, R. A. (2005b). Logistic regression *Advanced and Multivariate Statistical Methods* (3rd ed., pp. 313-330). Los Angeles, CA: Pyrczak Publishing.

Meyer, C. (2013). Applying evidence-based practice to the education debate. *NATA News, 25*, 8-9.

Michel, A. K., Colston, M. A., & Tanner, J. L. (2011). *Analysis of college volleyball player characteristics relevant to injury risk*. Paper presented at the Graduate Research Day, University of Tennessee at Chattanooga, Chattanooga, TN.

Middlemas, D. A., Manning, J. M., Gazzillo, L. M., & Young, J. (2001). Predicting performance on the National Athletic Trainers' Association Board of Certification examination from grade point average and number of clinical hours. *Journal of Athletic Training, 36*(2), 136-140.

Mitchell, K. J. (1990). Traditional predictors of performance in medical school. *Academic Medicine, 65*(3), 149-158.

Morris, J., & Farmer, A. (1999). The predictive strength of entry grades and biographical factors on the academic and clinical performance of physiotherapy students. *Physiotherapy Theory & Practice, 15*(3), 165.

Morrison, T., & Morrison, M. (1995). A meta-analytic assessment of the predictive validity of the quantitative and verbal components of the Graduate Record Examination with graduate grade point average representing the criterion of graduate success. *Educational and Psychological Measurement 55* (2), 309-316. doi: 10.1177/0013164495055002015

Morrison, T. M., Bruce, S. L., & Wilkerson, G. B. (2012). *Pre-participation injury risk status as a predictor of medical expenditures for college football players*. Paper presented at the Orthopedic Research Grand Rounds, Chattanooga, TN.

Munro, B. H. (1985). Predicting success in graduate clinical specialty programs. . *Nursing Research, 34*(1), 54-57. doi: 10.1097/00006199-198501000-198500011.

Munro, B. H. (2005a). Logistic regression *Statistical Methods for Health Care Research* (5th ed., pp. 301-320). Philadelphia, PA: Lippincott Williams & Wilkins.

Munro, B. H. (2005b). *Statistical methods for health care research* (5th ed.). Philadelphia, PA: Lippencott Williams & Wilkins.

National Athletic Trainers' Association. (2000). Guideline technical standards for entry-level athletic training education Retrieved June 12, 2012, from http://www.nata.org/education/educational-programs/technical-standards

National Athletic Trainers' Association. (2011a). National Athletic Trainers' Association history Retrieved September 12, 2011, from http://www.nata.org/nata-history

National Athletic Trainers' Association. (2011b). Professional education programs Retrieved September 17, 2011, from http://www.nata.org/ProfessionalEduPrgms

National Athletic Trainers Association Board of Certification. (2005). 2004 annual report for the National Athletic Trainers Association Board of Certification, Inc. Princeton, NJ.

National Athletic Trainers Association Board of Certification. (2006). 2005 annual report for the National Athletic Trainers' Association Board of Certification, Inc. Princeton, NJ.

National Athletic Trainers' Association. (2011). Athletic training education competencies (5th ed.). Dallas, TX.

National Athletic Trainers' Association Board of Certification. (2003). 2002 annual report Princeton, NJ.

National Athletic Trainers' Association Board of Certification. (2004). 2003 annual report for the National Athletic Trainers Association Board of Certification, Inc. Princeton, NJ.

National Athletic Trainers' Association Board of Certification Inc. (1997). Report on the 1996 certification examination. Princeton, NJ.

National Athletic Trainers' Association Board of Certification Inc. (1999). Annual report for the 1998 testing year. Princeton, NJ.

National Athletic Trainers' Association Board of Certification Inc. (2000). Annual report for the 1999 testing year. Princeton, NJ.

National Athletic Trainers' Association Board of Certification Inc. (2002). Annual examination report: 2001. Princeton, NJ.

National Athletic Trainers' Association Board of Certification Inc. (2006). BOC standards of professional practice Retrieved June 10, 2012, from http://www.bocatc.org/images/stories/multiple_references/standardsprofessionalpractice.pdf

National Athletic Trainers' Association Board of Certification Inc. (2007). 2006 annual report for the National Athletic Trainers' Association Board of Certification. Inc. Princeton, NJ.

National Board for Certification in Occupational Therapy. (2009). Welcome Retrieved June 10, 2012, from http://www.nbcot.org/

Newton, S. E., & Moore, G. (2007). Undergraduate grade point average and Graduate Record Examination scores: The experience of one graduate nursing program. *Nursing Education Perspectives, 28*(6), 327-331.

Oreopoulos, P., & Petronijevic, U. (2013). Making college worth it: A review of the returns to higher education. *The Future of Children, 23*(1), 41-65.

Pasco, A. H. (2009). Should graduate students publish? *Journal of Scholarly Publishing, 40*(3), 231-240.

Payton, O. D. (1997). A meta-analysis of the literature on admissions criteria as predictions of academic performance in physical therapy education in the United States and Canada: 1983 through 1994. *Physiotherapy Canada, 49*(2), 97.

Peng, C. Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The Journal of Educational Research, 96*(1), 3-14.

Peng, C. Y. J., & So, T.-S. H. (2002). Logistic regression analysis and reporting: A primer. *Understanding Statistics, 1*(1), 37-70.

Perdew, P. R. (2001). Developmental education and Alfred Binet: The original purpose of standardized testing. In J. L. Higbee, D. B. Lundell, I. M. Duranczyk & D. Banerjee-Stevens (Eds.), *2001: A Developmental Odyssey*. Warrensburg, MO: National Association for Developmental Education.

Peters, T. J. (2008). Multifarious terminology: Multivariable or multivariate? Univariable or univariate? *Paediatric and Perinatal Epidemiology, 22*(6), 506.

Phillips, B., Ball, C., Sackett, D., Badenoch, D., Straus, S., Haynes, B., & Dawes, M. (2009). Oxford Centre for Evidence-based Medicine levels of evidence *British Journal Urinary International, 103*(8), 1147. doi: 10.1111/j.1464-410X.2009.08556.x

Pickard, J. V. (2003). *An examination of the relationship between the mentorship of student athletic trainers and their outcome on the National Athletic Trainers' Association certification examination.* (Doctor of Education Dissertation), Sam Houston State University, Huntsville, TX.

Pitney, W. A. (2012). Requiring professional athletic training programs at the post-baccalaureate level: Considerations and concerns. *Athletic Training Education Journal, 7*(1), 4-5. doi: 10.5608/070104

Platt, L. S., Sammarone-Turocy, P., & McGlumphy, B. E. (2001). Preadmission criteria as predictors of academic success in entry-level athletic training and other allied health educational programs. *Journal of Athletic Training, 36*(2), 141-144.

Podichetty, V. K., & Morisue, H. (2009). Prediction rules in cervical spine injury. *British Medical Journal, 339*.

Portney, L. G., & Watkins, M. P. (2000). *Foundations of clinical research: Applications to practice* (2^{nd} ed.). Upper Saddle River, NJ: Prentice Hall.

Prentice, W. E. (2013, August/September). Is a transition to the Entry-level Master's degree really the best choice for the profession? *NATA News, 25,* 10-11.

Prieto-Marañón, P., Aguerri, M. E., Galibert, M. S., & Attorresi, H. F. (2012). Detection of differential item functioning: Using decision rules based on the Mantel-Haenszel procedure and Breslow-Day tests. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 8*(2), 63-70. doi: 10.1027/1614-2241/a000038

Reboldi, G., Angeli, F., & Verdecchia, P. (2013). Multivariable analysis in cerebrovascular research: Practical notes for the clinician. [Review]. *Cerebrovascular Diseases, 35*(2), 187-193.

Reinecke, M., & Wilkerson, G. B. (2012). *Risk factor for lateral ankle sprains in Division I-FCS football players*. Paper presented at the Southeastern Athletic Trainers’ Association 3rd Biennial Athletic Training Educators’ Conference,, Atlanta, GA.

Rhodes, M. L., Bullough, B., & Fulton, J. (1994). The Graduate Record Examination as an admission requirement for the graduate nursing program. *Journal of Professional Nursing, 10*(5), 289-296.

Richardson, C. A., Snijders, C. J., Hides, J. A., Damen, L., Pas, M. S., & Storm, J. (2002). The relation between the transversus abdominis muscles, sacroiliac joint mechanics, and low back pain. *Spine*, *27*(4), 399-405.

Rigney, E. R., & Bruce, S. L. (2010). *Prediction of knee injury in female collegiate basketball players*. Paper presented at the Celebration of Graduate Student Scholarship, University of Tennessee at Chattanooga, Chattanooga, TN.

Risberg, T. F. (2010). National standards and tests: The worst solution to america's educational problems-except for all the others. *George Washington Law Review, 79*(3), 890-925.

Rojstaczer, S. (2009, March 29). Grade inflation at American colleges and universities Retrieved June 11, 2012, from http://gradeinflation.com/

Rosin, A., & Sinopoli, M. (1999). Impact of the Ottawa ankle rules in a U.S. Army troop medical clinic in South Korea. *Military Medicine*, *164*(11), 793.

Rothman, K. J., Greenland, S., & Lash, T. L. (2008). *Modern epidemiology*. Philadelphia: Lippincott Williams & Wilkins.

Ruopp, M. D., Perkins, N. J., Whitcomb, B. W., & Schisterman, E. F. (2008). Youden index and optimal cut-point estimated from observations affected by a lower limit of detection. *Biometrical Journal*, *50*(3), 419-430. doi: 10.1002/bimj.200710415

Sackett, D. L. (1997). Evidence-based medicine. *Seminars in Perinatology*, *21*(1), 3-5.

Salvatori, P. (2001). Reliability and validity of admissions tools used to select students for the health professions *Advances in Health Sciences Education*, *6*(2), 159-175. doi: 110.1023/A:1011489618208.

Saupe, J. L., & Eimers, M. T. (2012). *Alternative estimates of the reliability of college grade point averages*. Paper presented at the Annual Forum of the Association for Institutional Research, New Orleans, LA. http://ir.missouri.edu/reports-presentations/AIR_Version_AlternativeEstimatesoftheReliability%20of%20College%20GPA_05-25-12

Shiyko, M. P., & Pappas, E. (2009). Validation of pre-admission requirements in a Doctor of Physical Therapy program with a large representation of minority students. *Journal of Physical Therapy Education*, *23*(2), 29-36.

Silver, B., & Hodgson, C. S. (1997). Evaluating GPAs and MCAT scores as predictors of NBME I and clerkship performances based on students' data from one undergraduate institution. *Academic Medicine*, *72*(5), 394-396.

Silver, N. (2012). *The signal and the noise: Why so many predictions fail-but some don't*. New York, NY: Penguin Press.

Sime, A. M., Corcoran, S. A., & Libera, M. B. (1983). Predicting success in graduate education. *Journal of Nursing Education, 22*(1), 7.

Singleton Jr., R., & Smith, E. R. (1978). Does grade inflation decrease the reliabiltiy of grades. *Journal of Educational Measurement, 15*(1), 37-41.

Snider, V. K., MacLean IV, D., & Wilkerson, G. (2013). *Pre-participation injury risk assessment and the effectiveness of programs for risk reduction in female collegiate athletes*. Paper presented at the Graduate Research Day, University of Tenneesee at Chattanooga, Chattanooga, TN.

Springer, B. A., Arciero, R. A., Tenuta, J. J., & Taylor, D. C. (2000). A prospective study of modified Ottawa ankle rules in a military population: Interobserver agreement between physical therapists and orthopaedic surgeons. *American Journal of Sports Medicine, 28*(6), 864.

Stanley, J. L., & Bruce, S. L. (2009). *A clinical prediction rule for overuse injuries of the upper extremity in college softball players*. Paper presented at the Celebration of Graduate Student Scholarship, University of Tennessee at Chattanooga, Chattanooga, TN.

Starkey, C., & Henderson, J. (1995). Performance on the athletic training certification examination based on candidates' routes to eligibility. *Journal of Athletic Training, 30*(1), 59-62.

Steves, R., & Hootman, J. (2004). Evidence-based medicine: What is it and how does it apply to athletic training? *Journal of Athletic Training, 39*, 83-87.

Steyerberg, E., & Harrell, J., Frank E. . (2003). Statistical models for prognostication. In M. B. Max & J. Lynn (Eds.), Interactive textbook of clinical symptom research. Bethesda, MD: National Institutes of Health. Retrieved from http://painconsortium.nih.gov/symptomresearch/chapter_8/index.htm.

Stiell, I. (1996). Ottawa ankle rules. *Canadian Family Physician, 42*, 478-480.

Stiell, I., Greenberg, G., McKnight, R., Nair, R., McDowell, I., & Worthington, J. (1992). A study to develop clinical decision rules for the use of radiography in acute ankle injuries. *Annals of Emergency Medicine, 21*(4), 384-390.

Straus, S. E., Richardson, W. S., Glasziou, P., & Haynes, R. B. (2005). *Evidence-based medicine: How to practice and teach EBM* (3rd ed.). London, England: Churchill Livingstone.

Stricker, G., & Huber, J. T. (1967). The Graduate Record Examination and undergraduate grades as predictors of success in graduate school. *The Journal of Educational Research, 60*(10), 466-468.

Sutlive, T. G., Lopez, H. P., Schnitker, D. E., Yawn, S. E., Halle, R. J., Mansfield, L. T., . . . Childs, J. D. (2008). Development of a clinical prediction rule for diagnosing hip osteoarthritis in individuals with unilateral hip pain. *Journal of Orthopaedic & Sports Physical Therapy*, *38*(9), 542-550.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson Education, Inc.

Templeton, M. S., Burcham, A., & Franck, L. (1994). Predictive study of physical therapy admission variables. *Journal of Allied Health*, *23*(2), 79.

Testing is Easy. (n.d.). What is the purpose of a standardized test? Retrieved January 16, 2013, from http://www.testingiseasy.com/standardized-test-purpose/

Teyhen, D. S., Flynn, T. W., Childs, J. D., & Abraham, L. D. (2007). Arthrokinematics in a subgroup of patients likely to benefit from a lumbar stabilization exercise program. *Physical Therapy*, *87*(3), 313-325.

Thacker, A. J., & Williams, R. E. (1974). The relationship of the graduate record examination to grade point average and success in graduate school. *Educational and Psychological Measurement*, *34* (4), 939-944. doi: 10.1177/001316447403400425

The Carnegie Foundation for the Advancement of Teaching. (2010). A classification of institutions of higher education: 2010 Edition. *The Carnegie Foundation for the Advancement of Teaching* Retrieved July 13, 2013, from http://classifications.carnegiefoundation.org/

The Faculty Committee on Grading. (2005). Grading at Princeton: Philosophy, strategy, practice Retrieved June 11, 2012, from http://www.princeton.edu/odoc/docs/Grading_at_Princeton.pdf

The Free Dictionary by Farlex. (2000a, 2009). competence. *The American Heritage® Dictionary of the English Language* 4th ed. Retrieved October 15, 2013, from http://www.thefreedictionary.com/competence

The Free Dictionary by Farlex. (2000b, 2009). variable. *The American Heritage® Dictionary of the English Language* 4th ed. Retrieved December 13, 2013, from http://www.thefreedictionary.com/variate

The Free Dictionary by Farlex. (2000c, 2009). variate. *The American Heritage® Dictionary of the English Language* 4th ed. Retrieved December 13, 2013, from http://www.thefreedictionary.com/variate

The Free Dictionary by Farlex. (2010). proficient. *Random House Kernerman Webster's College Dictionary* 4th ed. Retrieved October 15, 2013, from http://www.thefreedictionary.com/proficient

Tsai, A. C. (2013). Achieving consensus on terminology describing multivariable analyses. *American Journal of Public Health, 103*(6), e1-e1. doi: 10.2105/ajph.2013.301234

Tucker, K. M., Mullis, M. A., Wilkerson, G. B., & Bruce, S. L. (2013). *Development of a prediction model for identification of high-cost sports injury cases*. Paper presented at the Graduate Research Day, University of Tennessee at Chattanooga, Chattanooga, TN.

Turocy, P. S. (2002). Overview of athletic training education research publications. *Journal of Athletic Training, 37*(4 suppl), S-162.

Turocy, P. S., Comfort, R. E., Perrin, D. H., & Gieck, J. H. (2000). Clinical experiences are not predictive of outcomes on the NATABOC examination. *Journal of Athletic Training, 35*(1), 70-75.

United States Medical Licensing Examination. (2012). What is USMLE? Retrieved June 10, 2012, from http://www.usmle.org/

Utzman, R. R., Riddle, D. L., & Jewell, D. V. (2007a). Use of demographic and quantitative admissions data to predict academic difficulty among professional physical therapist students. *Physical Therapy*, *87*(9), 1164-1180.

Utzman, R. R., Riddle, D. L., & Jewell, D. V. (2007b). Use of demographic and quantitative admissions data to predict performance on the National Physical Therapy Examination. *Physical Therapy*, *87*(9), 1181-1193.

Vallverdú, J. (2008). The false dilemma: Bayesian vs. Frequentist. *arXiv preprint arXiv:0804.0486*.

Vendrely, A. M. (2007). An investigation of the relationships among academic performance, clinical performance, critical thinking, and success on the physical therapy licensure examination. *Journal of Allied Health, 36*(2), e108-123.

Verhagen, E., & Van Mechelen, W. (2009). *Sports injury research* (Vol. 11): Oxford University Press.

Vincent, W. J., & Weir, J. P. (2012). *Statistics in kinesiology* (4th ed.). Champaign, IL: Human Kinetics Publishers.

Wainner, R. S., Fritz, J. M., Irrgang, J. J., Delitto, A., Allison, S., & Boninger, M. L. (2005). Development of a clinical prediction rule for the diagnosis of carpal tunnel syndrome. *Archives of Physical Medicine and Rehabilitation, 86*, 609-618.

Warren, J. R. (1971). College grading practices: An overview (pp. 1-29). ERIC Clearinghouse on Higher Education: The George Washington University.

Wasson, J. H., Sox, H. C., Neff, R. K., & Goldman, L. (1985). Clinical prediction rules: Applications and methodological standards. *New England Journal of Medicine*, *313*(13), 793-799.

Weidner, T. G., & Henning, J. M. (2002). Historical perspective of athletic training clinical education. *Journal of Athletic Training, 37*(4 suppl), S-222 - S-228.

Werts, C., Linn, R. L., & Jöreskog, K. G. (1978). Reliability of college grades from longitudinal data. *Educational and Psychological Measurement, 38*, 89-95.

Westphalen, S. W., & McLean, J., Lindsy (1978). Seven years of certification by the NATA. *Athletic Training - The Journal of the National Athletic Trainers' Association, 13*(2), 86, 88, 91.

Wilkerson, G. B. (2011, October 18). [Discussion of dichotomizing independent variables verses using continuous variables in a logistic regression].

Wilkerson, G. B. (2012). *Prediction of injury risk in college football players*. Paper presented at the 37th Annual Southeast Athletic Trainers' Association Clinical Symposia & Members Meeting Atlanta, GA.

Wilkerson, G. B., Bullard, J. T., & Bartal, D. W. (2010). Identification of cardiometabolic risk among collegiate football players. *Journal of Athletic Training, 45*(1), 67-74. doi: 10.4085/1062-6050-45.1.67

Wilkerson, G. B., Colston, M. A., & Bogdanowicz, B. T. (2006). Distinctions between athletic training education programs at the undergraduate and graduate levels. *Athletic Training Education Journal*, *1*(2), 38-40.

Wilkerson, G. B., Giles, J. L., & Seibel, D. K. (2012). Prediction of core and lower extremity strains and sprains in collegiate football players: A preliminary study. *Journal of Athletic Training*, *47*(3), 264-272.

Williams, J. D., Harlow, S. D., & Stable, D. G. (1970). A longitudinal study examining prediction of doctoral success: Grade point average as criterion, or graduation vs. non-graduation as criterion *The Journal of Educational Research, 64*(4), 161-164.

Williams, R. B., & Hadfield, O. D. (2003). Attributes of curriculum athletic training programs related to the passing rate of first-time certification examinees. *Journal of Allied Health, 32*(4), 240-245.

Willingham, W. W. (1972). *Predicting success in graduate education.* Paper presented at the Council of Graduate School.

Wilson, T. (1999). A student selection method and predictors of success in a graduate nursing program. *Journal of Nursing Education, 38*(4), 183-187.

Wolk, R. A. (2009). Why we're still 'at risk': The legacy of five faulty assumptions. *Education Week, 28*(29), 30, 36.

Yealy, D. M., & Auble, T. E. (2003). Choosing between clinical prediction rules. *New England Journal of Medicine, 349*(26), 2553.

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer, 3*(1), 32-35. doi: 10.1002/1097-0142(1950)3:1<32::aid-cncr2820030106>3.0.co;2-3

Yuen, M. (2001). The Ottawa ankle rules in children. *Emergency Medicine Journal, 18*(6), 466.

Zabell, S. (1989). R. A. Fisher on the history of inverse probability. *Statistical Science, 4*(3), 247-256.

Zaglaniczny, K. L. (1992). Factors which predict performance on the National Certification Examination for Nurse Anesthetists. *American Association of Nurse Anesthetists Journal, 60*(6), 533-540.

Zipp, G. P., Ruscingno, G., & Olson, V. (2010). Admission variables and academic success in the first year of the professional phase in a doctor of physical therapy program. *Journal of Allied Health, 39*(3), 138-142.

APPENDIX A

UNIVARIABLE ANALYSES AND 2 x 2 CROSS TABULATION TABLES

FOR POTENTIAL PREDICTORS RELATED TO FIRST-ATTEMPT

PASS – YES OR NO ON THE BOC EXAM

APPENDIX A

Univariable analysis results for each of the potential predictors related to first-attempt pass – Yes or No, on the BOC exam are provided in Figures A.1 to A.7 and Tables A.1 to A.7.

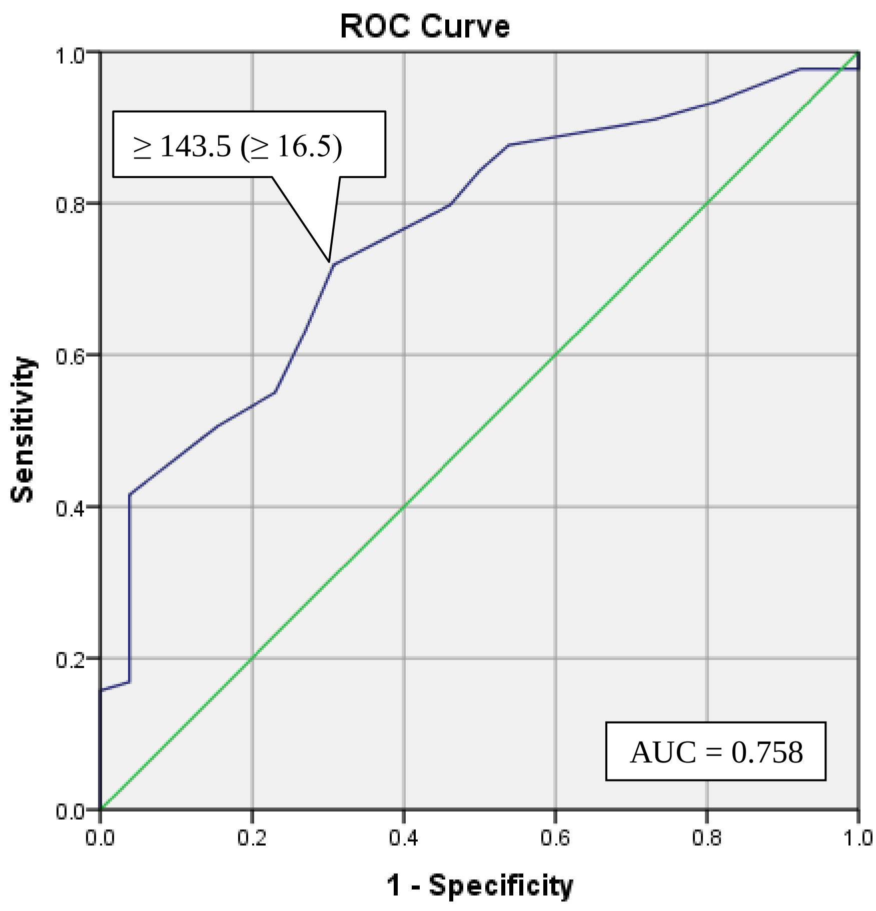

Figure A.1 ROC curve with identification of the optimum cut-point for GREq (PR) for prediction of first-attempt BOC exam success

Table A.1 GREq (PR) score for prediction of first-attempt BOC exam success

<table>
<tr><td rowspan="2"></td><td colspan="2">1st Attempt Pass on the BOC exam</td></tr>
<tr><td>Yes</td><td>No</td></tr>
<tr><td>≥ 143.5 (16.5)</td><td>64</td><td>8</td></tr>
<tr><td>< 143.5 (16.5)</td><td>25</td><td>18</td></tr>
<tr><td colspan="3">Fisher’s Exact Test (one-sided) p < 0.001</td></tr>
<tr><td colspan="2">Sn = 0.72 (95% CI: 0.62 – 0.80)</td><td>Sp = 0.69 (95% CI: 0.500 – 0.84)</td></tr>
<tr><td colspan="3">Youden’s Index = 0.411</td></tr>
<tr><td colspan="2">OR = 5.76 (95% CI: 2.22 – 14.93)</td><td>RFS = 1.53 (95% CI: 1.25 – 1.87)</td></tr>
</table>

A student in the GATP who had a GREq score of ≥ 143.5 (PR ≥ 16.5) , had 5.76 times greater odds of passing the BOC exam on their first-attempt than the odds for someone who had a GREq score of < 143.5 (PR < 16.5).

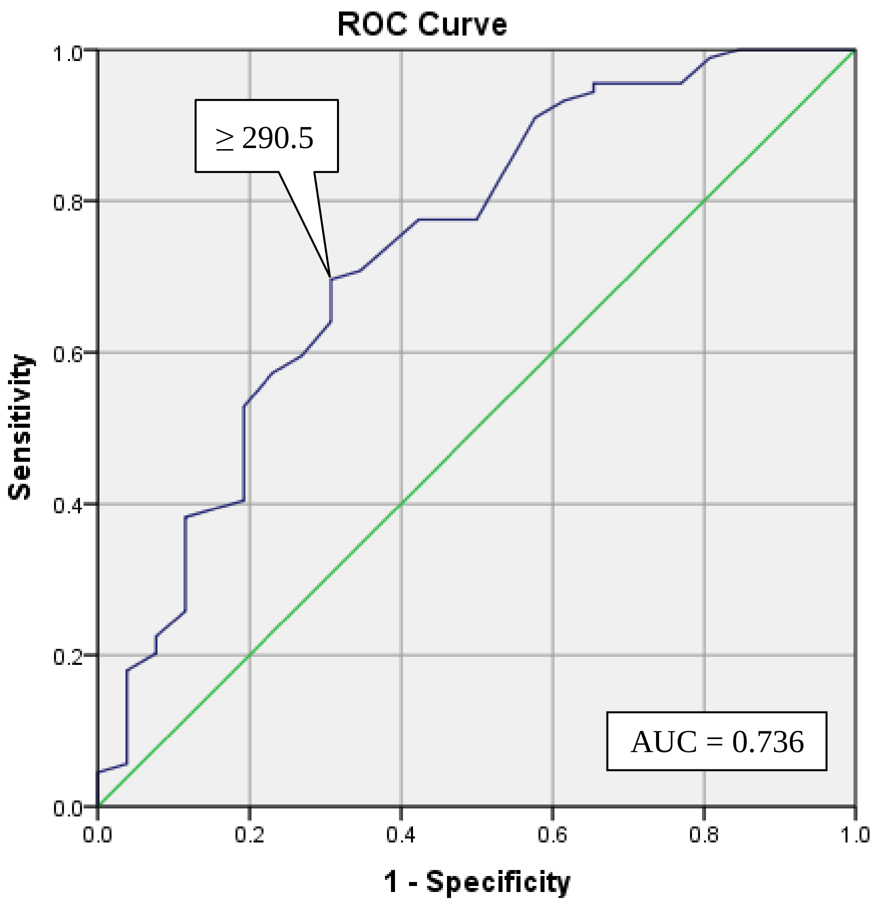

Figure A.2 ROC curve with identification of optimum cut-point for GRE – Composite score for prediction of first-attempt BOC exam success

Table A.2 GRE – Composite score for prediction of first-attempt BOC exam success

<table>
<tr><td rowspan="2"></td><td colspan="2">1st Attempt Pass on the BOC exam</td></tr>
<tr><td>Yes</td><td>No</td></tr>
<tr><td>≥ 290.5</td><td>55</td><td>10</td></tr>
<tr><td>< 290.5</td><td>34</td><td>19</td></tr>
<tr><td colspan="3">Fisher’s Exact Test (one-sided) p = 0.001</td></tr>
<tr><td>Sn = 0.70 (95% CI: 0.60 – 0.78)</td><td colspan="2">Sp = 0.69 (95% CI: 0.50 – 0.84)</td></tr>
<tr><td colspan="3">Youden’s Index = 0.389</td></tr>
<tr><td>OR = 5.17 (95% CI: 2.00 – 13.33)</td><td colspan="2">RFS = 1.48 (95% CI: 1.20 – 1.81)</td></tr>
</table>

A student in the GATP who had GRE – Composite score of 290.5 or greater, had 5.17 times greater odds of passing the BOC exam on their first-attempt than the odds for someone who had a GRE - Composite score of less than 290.5.

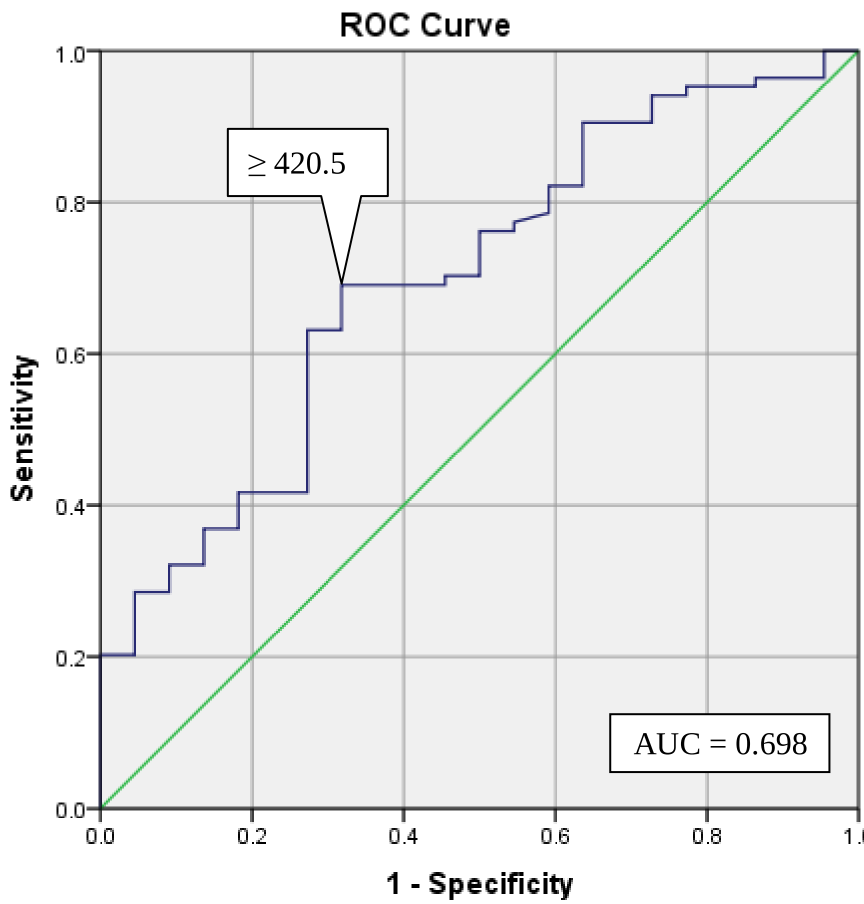

Figure A.3 ROC curve with identification of the optimum cut-point for Biderman's Formula Score for prediction of first-attempt BOC exam success

Table A.3 Biderman's Formula Score for prediction of first-attempt BOC exam success

	1st Attempt Pass on the BOC exam	
	Yes	**No**
≥ 420.5	58	7
< 420.5	26	15
Fisher's Exact Test (one-sided) $p = 0.003$		
Sn = 0.69 (95% CI: 0.59 – 0.78)	Sp = 0.68 (95% CI: 0.47 – 0.84)	
Youden's Index = 0.372		
OR = 4.78 (95% CI: 1.74 – 13.12)	RFS = 1.41 (95% CI: 1.15 – 1.73)	

A student in the GATP, who had a Biderman's Formula Score of 420.5 or greater, had 4.78 times greater odds of passing the BOC exam on their first-attempt than the odds for someone who had a Biderman's Formula Score of less than 420.5.

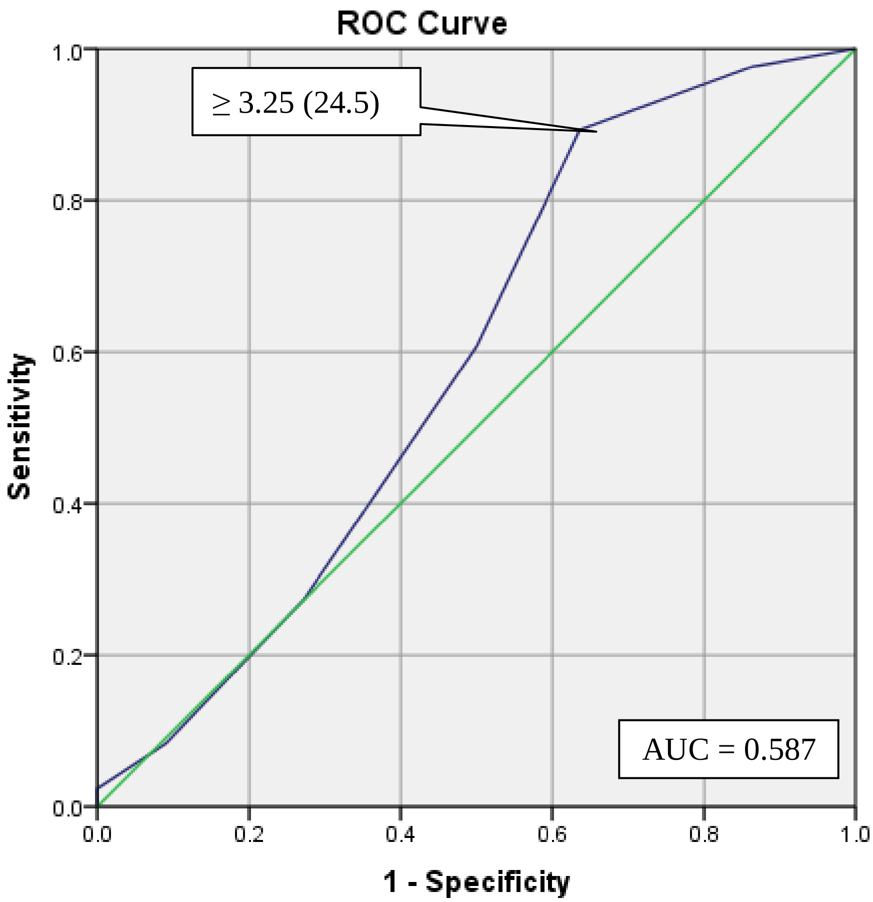

Figure A.4 ROC curve with identification of the optimum cut-point for GREwr (PR) for prediction of first-attempt BOC exam success

Table A.4 GREwr (PR) score for prediction of first-attempt BOC exam success

	1st Attempt Pass on the BOC exam	
	Yes	**No**
≥ 3.25 (24.5)	75	14
< 3.25 (24.5)	9	8
Fisher's Exact Test (one-sided) $p = 0.007$		
Sn = 0.89 (95% CI: 0.81 – 0.94)	Sp = 0.36 (95% CI: 0.20 – 0.57)	
Youden's Index = 0.257		
OR = 4.76 (95% CI: 1.57 – 14.45)	RFS = 1.59 (95% CI: 1.30 – 1.95)	

A student in the GATP who had a GREwr score of ≥ 3.25, (PR ≥ 24.5), had 4.76 times greater odds of passing the BOC exam on their first-attempt than the odds for someone who had a GREwr score of < 3.25 (PR < 24.5).

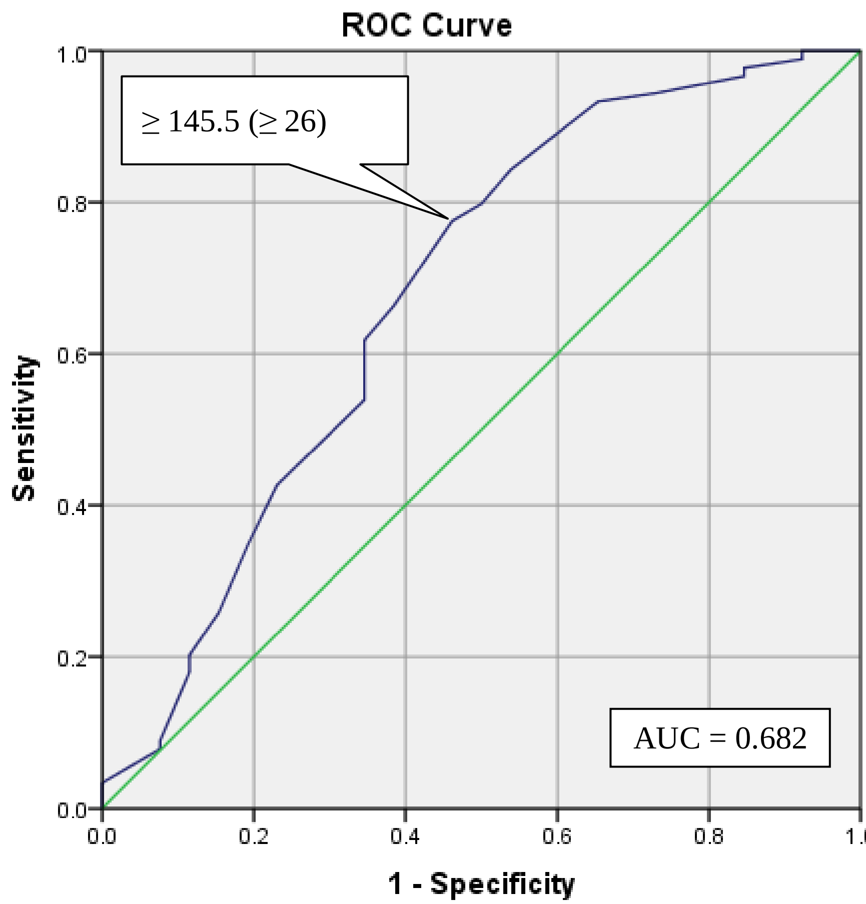

Figure A.5 ROC curve with identification of the optimum cut-point for GREv (PR) for prediction of first-attempt BOC exam success

Table A.5 GREv (PR) score for prediction of first-attempt BOC exam success

	1st Attempt Pass on the BOC exam	
	Yes	**No**
≥ 145.5 (26)	69	12
< 145.5 (26)	20	14
Fisher's Exact Test (one-sided) *p* = 0.005		
Sn = 0.78 (95% CI: 0.68 – 0.85)	Sp = 0.54 (95% CI: 0.36 – 0.71)	
Youden's Index = 0.538		
OR = 4.25 (95% CI: 1.61 – 10.11)	RFS = 1.45 (95% CI: 1.18 – 1.78)	

A student in the GATP who had a GREv score of ≥ 145.5, (PR ≥ 26) had 4.25 times greater odds of passing the BOC exam on their first-attempt than the odds for someone who had a GREv score of < 145.5 (PR < 26)

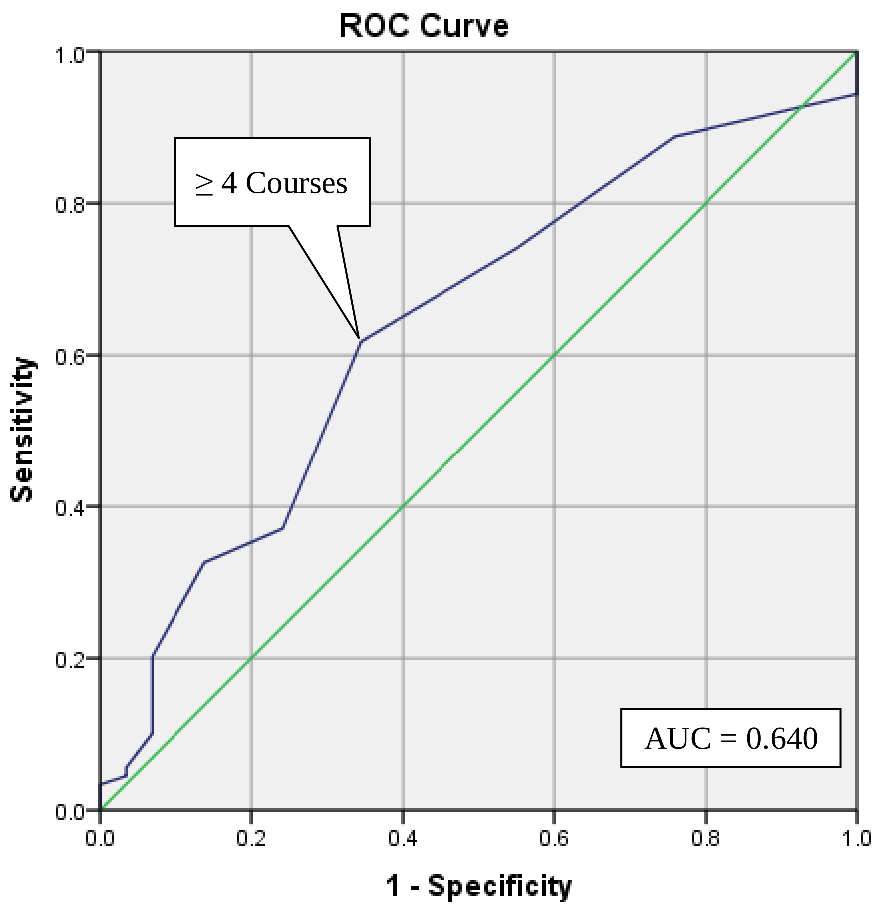

Figure A.6 ROC curve with identification of the optimum cut-point for the number of advanced math, science or athletic training courses for prediction of first-attempt BOC exam success

Table A.6 Number of advanced math, science or athletic training courses for prediction of first-attempt BOC exam success

	1st Attempt Pass on the BOC exam	
	Yes	**No**
≥ 4 Courses	55	10
< 4 Courses	34	19
Fisher's Exact Test (one-sided) p = 0.017		
Sn = 0.62 (95% CI: 0.51 – 0.71)	Sp = 0.66 (95% CI: 0.47 – 0.80)	
Youden's Index = 0.273		
OR = 3.07 (95% CI: 1.28 – 7.39)	RFS = 1.32 (95% CI: 1.08 – 1.62)	

A student in GATP who took four or more advanced math, science or athletic training courses had 3.07 times greater odds of passing the BOC exam on their first-attempt than the odds for someone who took less than four advanced math, science or athletic training courses.

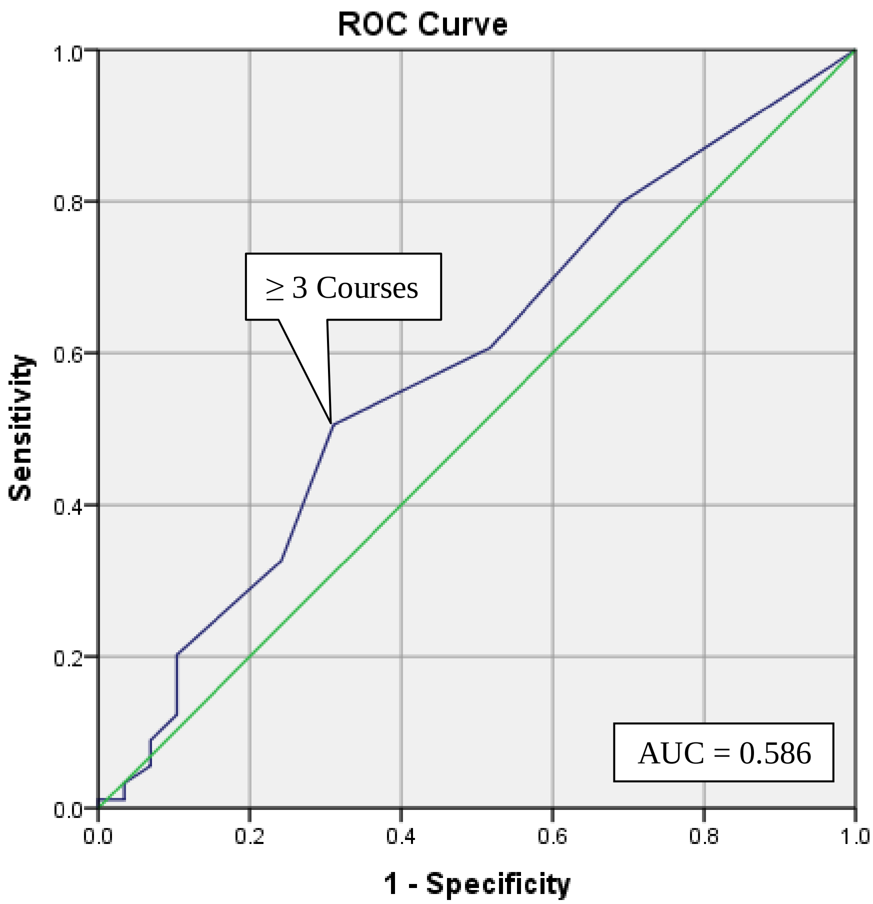

Figure A.7 ROC curve with identification of the optimum number of advanced math and science courses for prediction of first-attempt BOC exam success

Table A.7 Number of advanced math and science courses for prediction of first-attempt BOC exam success

	1st Attempt Pass on the BOC exam	
	Yes	**No**
≥ 3 Courses	45	9
< 3 Courses	44	20
Fisher's Exact Test (one-sided) p = 0.087		
Sn = 0.51 (95% CI: 0.40 – 0.61)	Sp = 0.70 (95% CI: 0.51 – 0.83)	
Youden's Index = 0.196		
OR = 2.27 (95% CI: 0.93 – 5.53)	RFS = 1.21 (95% CI: 0.99 – 1.49)	

A student in the GATP who took three or more advanced math and science courses had 2.27 times greater odds of passing the BOC exam on their first-attempt than the odds for someone who took less than three advanced math and science courses.

APPENDIX B

UNIVARIABLE ROC ANALYSES AND 2 x 2 CROSS TABULATION TABLES FOR POTENTIAL PREDICTORS RELATED TO ACADEMIC PROFILE OF UNDERGRADUATE INSTITUTIONS (APUI)

APPENDIX B

Univariable analysis results for each of the potential APUI-related predictors of gGPA at the end of the first-year are provided in Tables B.1 to B.8 and Figures B.1 and B.2

Table B.1 Percentile statistics for undergraduate institutions (N = 194) of students admitted to GATP

		Institution SAT mean/median (N = 110)	**Institution ACT mean/median (N = 121)**
Percentiles for all undergraduate institutions represented	50	1143.5	24.0
	75	1195.0	26.0
	80	1238.0	27.0

Table B.2 Summary of univariable analysis results for predictions of first-year gGPA (≥ 3.45) derived from APUI data for students admitted to GATP

Academic Profile of Undergraduate Institution	**Cut-point**	**Sn**	**1 - Sp**	**Sp**	**Youden's Index**	**AUC**	**OR**	**RFS**	**Fisher's Exact Test (one-sided) *p*-value**
Institution SAT mean/median ≥ 80th pctl	≥ 1238.0	0.28		0.97	–	–	11.58	1.46	0.002
Institution ACT mean/median	≥ 25.5	0.48	0.14	0.86	0.341	0.710	5.82	1.54	0.001
Institution ACT mean/median ≥ 75th pctl	≥ 26.0	0.48		0.86	–	–	5.82	1.54	0.001
Institution ACT mean/median ≥ 80th pctl	≥ 27.0	0.29		0.92	–	–	4.53	1.39	0.009
Institution SAT mean/median ≥ 75th pctl	≥ 1195	0.32		0.90	–	–	4.32	1.36	0.013
Institution SAT mean/median	≥ 1132.5	0.61	0.29	0.71	0.318	0.697	3.78	1.44	0.003

Table B.3 Institution SAT mean/median ≥ 80th pctl for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
≥ 1238	22	1
< 1238	57	30
Fisher's Exact Test (one-sided) *p* = 0.002		
Sn = 0.28 (95% CI: 0.19 – 0.39)	Sp = 0.97 (95% CI: 0.84 – 0.99)	
OR = 11.58 (95% CI: 1.49 – 90.14)	RFS = 1.46 (95% CI: 1.19 – 1.79)	

The OR of 11.58 (Fisher's Exact Test (one-sided) *p* = 0.002) for Institution SAT mean/median 80th pctl ≥ 1238 cut-point met the criterion for inclusion in a multi-variable analysis of potential predictors.

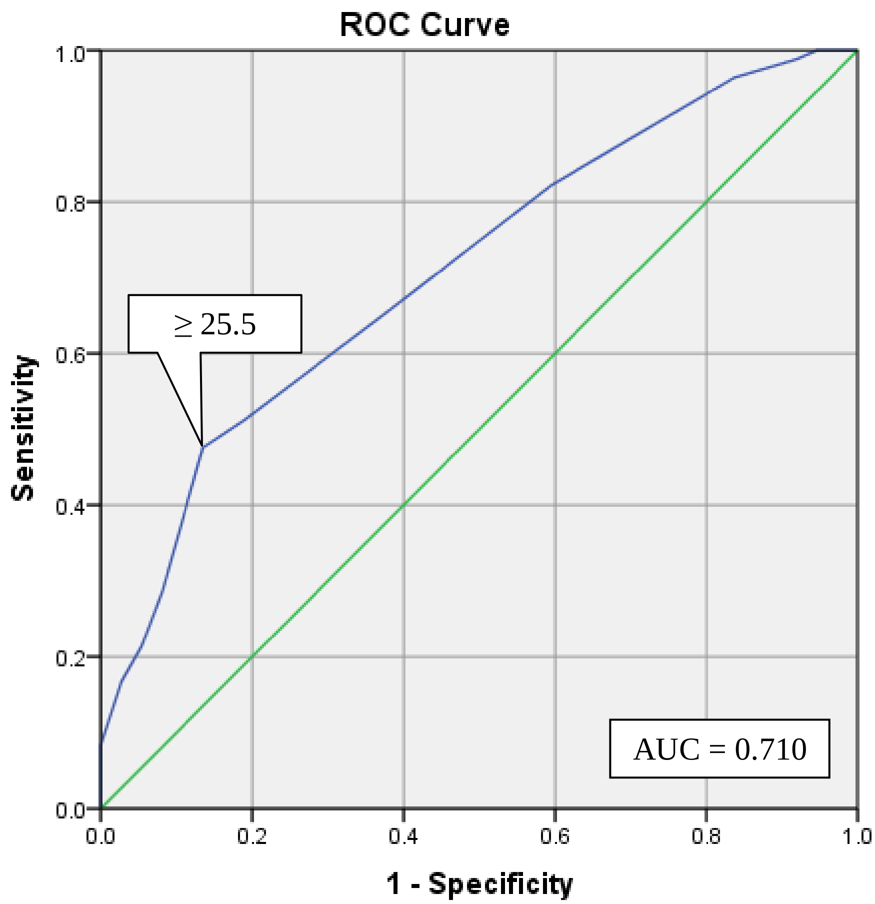

Figure B.1 ROC curve with identification of the optimum cut-point for Institutions' ACT mean/median for prediction of first-year gGPA (≥ 3.45)

Table B.4 Institution ACT mean/median for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
≥ 25.5	40	5
< 25.5	44	32
Fisher's Exact Test (one-sided) *p* < 0.001		
Sn = 0.48 (95% CI: 0.37 – 0.58)	Sp = 0.86 (95% CI: 0.72 – 0.94)	
Youden's Index = 0.341		
OR = 5.82 (95% CI: 2.07 – 16.38)	RFS = 1.54 (95% CI: 1.25 – 1.88)	

The OR of 5.82 (Fisher's Exact Test (one-sided) $p < 0.001$) for Institution ACT mean/median ≥ 25.5 cut-point met the criterion for inclusion in a multi-variable analysis of potential predictors.

Table B.5 Institution ACT mean/median ≥ 75th pctl for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
≥ 26	40	5
< 26	44	32
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.48 (95% CI: 0.37 – 0.58)	Sp = 0.86 (95% CI: 0.72 – 0.94)	
OR = 5.82 (95% CI: 2.07 – 16.38)	RFS = 1.54 (95% CI: 1.25 – 1.88)	

The OR of 11.58 (Fisher's Exact Test (one-sided) $p = 0.002$) for Institution SAT mean/median 80th pctl ≥ 1238 cut-point met the criterion for inclusion in a multi-variable analysis of potential predictors.

Table B.6 Institution ACT mean/median ≥ 80th pctl for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
≥ 27	24	3
< 27	60	34
Fisher's Exact Test (one-sided) $p = 0.009$		
Sn = 0.29 (95% CI: 0.20 – 0.39)	Sp = 0.92 (95% CI: 0.79 – 0.97)	
OR = 4.53 (95% CI: 1.27 – 16.17)	RFS = 1.39 (95% CI: 1.14 – 1.71)	

The OR of 11.58 (Fisher's Exact Test (one-sided) p = 0.002) for Institution SAT mean/median 80th pctl ≥ 1238 cut-point met the criterion for inclusion in a multi-variable analysis of potential predictors.

Table B.7 Institution SAT mean/median ≥ 75th pctl for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
≥ 1195	25	3
< 1195	54	28
Fisher's Exact Test (one-sided) p = 0.13		
Sn = 0.32 (95% CI: 0.23 – 0.43)	Sp = 0.90 (95% CI: 0.75 – 0.97)	
OR = 4.32 (95% CI: 1.20 – 15.57)	RFS = 1.36 (95% CI: 1.11 – 1.66)	

The OR of 4.32 (Fisher's Exact Test (one-sided) p = 0.013) for Institution SAT mean/median ≥ 1195 cut-point met the criterion for inclusion in a multi-variable analysis of potential predictors.

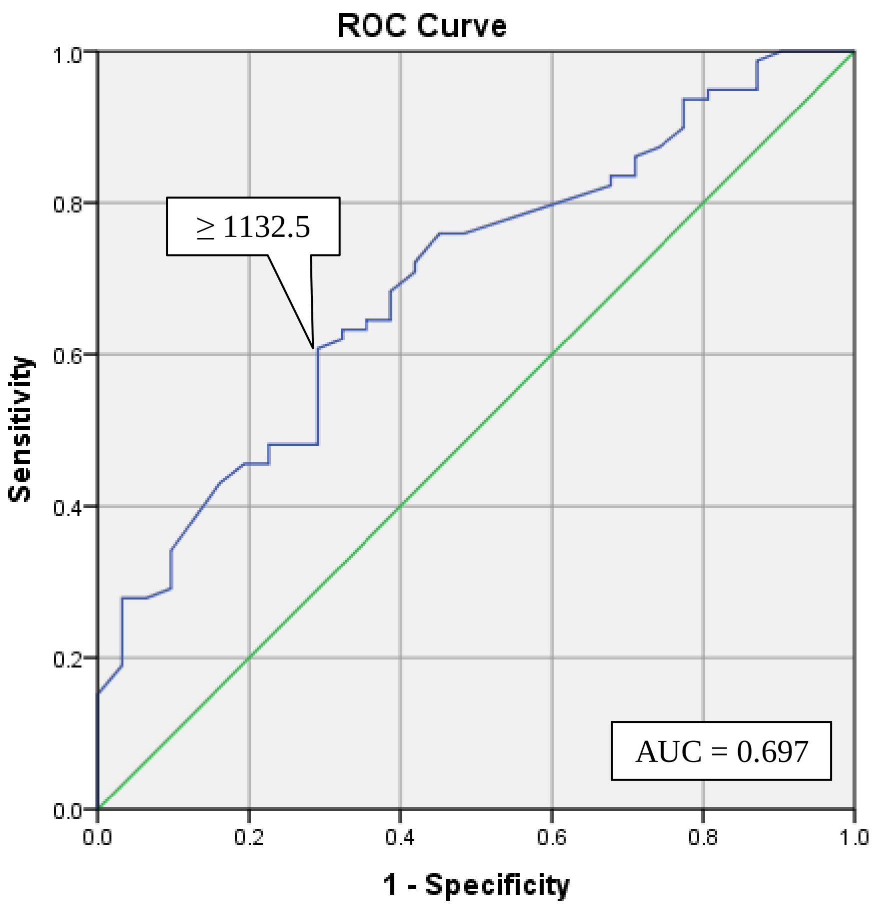

Figure B.2 ROC curve with identification of the optimum cut-point for Institution SAT mean/median for prediction of first-year gGPA (≥ 3.45)

Table B.8 Institution SAT mean/median for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
≥ 1132.5	48	9
< 1132.5	31	22
Fisher's Exact Test (one-sided) $p = 0.003$		
Sn = 0.61 (95% CI: 0.50 – 0.71)	Sp = 0.71 (95% CI: 0.53 – 0.89)	
Youden's Index = 0.318		
OR = 3.78 (95% CI: 1.54 – 9.29)	RFS = 1.44 (95% CI: 1.18 – 1.77)	

The OR of 3.78 (Fisher's Exact Test (one-sided) $p = 0.003$) for Institution SAT mean/median ≥ 1132.5 cut-point met the criterion for inclusion in a multi-variable analysis of potential predictors.

APPENDIX C

ACT/SAT MEAN/MEDIAN SCORES FOR UNDERGRADUATE COLLEGES AND UNIVERSITIES ATTENDED BY GATP STUDENTS

APPENDIX C

Table C.1 Undergraduate colleges and universities attended by GATP students with ACT and SAT mean/median scores

Undergraduate College or University for GATP Students	ACT mean/median	SAT mean/median
Alma College	24	1080
Auburn University	27	1180
Austin Peay State University	22	980
Belhaven University	22	1030
Bellarmine University	24	1080
Berry College	26	1145
Bethel College	25	1156
Brevard College	19	910
California Polytechnic State University	29	1311
California State University - Chico	22	1020
California State University - Northridge	19	925
California State University - Sacramento	20	945
Centre College	28	1265
Clemson University	28.5	1245
College of Charleston	25	1205
Colorado State University	24	1140
Cornell University	26	1238
Dartmouth College	32	1455
East Tennessee State University	22	995
Elmhurst College	24	1098
Emmanuel College	24	1105
Eureka College	23	1165
Fayetteville State University	18	860
Florida A&M University	20	950
Freed-Hardeman University	24	1060
Friends University	22	1030
Gannon University	23	1050
Georgia College & State University	26	1170

Undergraduate College or University for GATP Students	ACT mean/median	SAT mean/median
Georgia Institute of Technology	30	1355
Georgia State University	23	1093
Gettysburg College		1285
Gonzaga University		
Grand Valley State University	24	1110
Hendrix College	29	1225
Houston Baptist University	24	1102
Huntington College	22	924
Indiana University - Bloomington	27	1195
Indiana University - Purdue - University	22	1005
Jacksonville State University	22.5	970
James Madison University	26	1190
Kennesaw State University	22	1075
King University (TN)		
Lee University	23	1070
Lipscomb University	25	1125
Longwood University	22	1030
Louisiana State University	25	1142
Maryville College	23	1070
Milligan College	23	1080
Mississippi State University	24	1110
Mississippi University for Women	22	1000
Northeastern University	30	1370
Oakland University	22	1030
Oglethorpe University	25	1145
Pepperdine University	29	1300
Pfeifer College	20	950
Rutgers - Newark		1045
Rutgers University - New Brunswick	27	1220
San Francisco State University	22	995
Santa Clara University	29	1290

Undergraduate College or University for GATP Students	ACT mean/median	SAT mean/median
Siena College	25	1160
Sonoma State University	20.5	1050
South Carolina State University	16.5	820
Southeastern LA University	22	1030
Southern University and A & M College	19	910
St Cloud State University	22	1046
St. Joseph's College	24	1125
SUNY - Fredonia	24	1090
Taylor University	27	1145
Tennessee State University	19	900
Tennessee Tech University	23.5	
Texas A&M University	27	1215
Texas Tech University	24	1115
Trevecca Nazarene University	22	950
Union University	26	1195
University of Alabama		
University of Alabama - Huntsville	26	1145
University of California - Davis	29	1295
University of California - Santa Barbara	28	1243
University of Central Florida	27	1245
University of Central Missouri	22	1030
University of Connecticut	28	1230
University of Delaware	29	1300
University of Florida	28	1265
University of Georgia	30	1355
University of Illinois – Urbana - Champaign	30	1370
University of Kentucky	25	1150
University of Louisville	24	1120
University of Memphis	22	1020
University of Minnesota	27.5	1295
University of Mississippi	18	830

Undergraduate College or University for GATP Students	ACT mean/median	SAT mean/median
University of Nevada - Reno	23.5	1065
University of North Carolina - Chapel Hill	30	1305
University of North Carolina - Greensboro	22	1035
University of Oregon	24	1110
University of Pittsburgh	28	1205
University of Portland		
University of Puget Sound	28	1249
University of South Alabama	26	1166
University of South Carolina - Aiken	24	1100
University of Tennessee - Chattanooga	23	1060
University of Tennessee - Knoxville	26	1175
University of Texas - Pan American	20	970
University of Washington	27	1215
University of Wisconsin - Oshkosh	22	1030
University of Wisconsin - Whitewater	22	1020
University of West Georgia	20.5	980
Valdosta State College	21.5	1030
Valparaiso University	26	
Virginia Tech University	28	1250
Wartburg College		
Wayne State University	22	1030
Western Michigan University	22	1030
Western Washington University	25	1125
Wheaton College	30	1320
Xavier University (OH)	25.5	1100
Xavier University of Louisiana	22	990
Youngstown State	20	950

APPENDIX D

MULTIVARIABLE ANALYSES AND 2 x 2 CROSS TABULATION TABLES FOR POTENTIAL PAIRS OF PREDICTORS RELATED TO ACADEMIC PROFILE OF UNDERGRADUATE INSTITUTIONS (APUI)

APPENDIX D

Multivariable analysis results for each of the potential pairs of predictors related to APUI for gGPA at the end of the first-year are provided in Tables D.1 to D.9.

A summary of potential predictor variables derived from APUI data are presented in Table D.1, which lists them in order of OR magnitude.

Table D.1 Summary of the best combination of reported Institution ACT and SAT scores to define high versus low AUPI

Combination of Institution ACT and SAT scores	**Sn**	**Sp**	**OR**	**RFS**	**Fisher's Exact Test (one-sided)**
Either ACT ≥ 25.5 or SAT 80th pctl ≥ 1238	0.47	0.88	6.51	1.56	0.001
Either ACT mean/median ≥ 25.5 or SAT 75th pctl ≥ 1195	0.47	0.88	6.51	1.56	0.001
Either ACT 75th pctl ≥ 26 or SAT 75th pctl ≥ 1195	0.47	0.88	6.47	1.56	0.001
Either ACT 75th pctl ≥ 26 or SAT 80th pctl ≥ 1238	0.47	0.88	6.47	1.56	0.001
Either ACT 80th pctl ≥ 27 or SAT 80th pctl ≥ 1238	0.32	0.93	6.09	1.46	0.001
Either ACT ≥ 25.5 or SAT mean/median ≥ 1132.5	0.56	0.81	5.39	1.59	0.001
Either ACT 75th pctl ≥ 26 or SAT mean/median ≥ 1132.5	0.56	0.78	4.51	1.52	0.001
Either ACT 80th pctl ≥ 27 or SAT mean/median ≥ 1132.5	0.55	0.79	4.54	1.52	0.001

Table D.2 Either Institution ACT mean/median ≥ 25.5 or an Institution SAT mean/median 80th pctl ≥ 1238 for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
Either ACT ≥ 25.5 or SAT 80th pctl ≥ 1238	44	5
Neither ACT ≥ 25.5 nor SAT 80th pctl ≥ 1238	50	37
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.47 (95% CI: 0.37 – 0.57)	Sp = 0.88 (95% CI: 0.75 – 0.95)	
OR = 6.51 (95% CI: 2.35 – 18.02)	RFS = 1.56 (95% CI: 1.28 – 1.92)	

The OR of 6.51 (Fisher's Exact Test (one-sided) $p < 0.001$) for Either Institution ACT mean/median ≥ 25.5 or an Institution SAT mean/median 80th pctl ≥ 1238 cut-point met the criterion for inclusion in a multi-variable analysis of potential predictors.

Table D.3 Either Institution ACT mean/median ≥ 25.5 or Institution SAT mean/median 75th pctl ≥ 1195 for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
Either ACT mean/median ≥ 25.5 or SAT 75th pctl ≥ 1195	44	5
Neither ACT mean/median ≥ 25.5 nor SAT 75th pctl ≥ 1195	50	37
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.47 (95% CI: 0.37 – 0.57)	Sp = 0.88 (95% CI: 0.75 – 0.95)	
OR = 6.51 (95% CI: 2.35 – 18.02)	RFS = 1.56 (95% CI: 1.28 – 1.92)	

The OR of 6.51 (Fisher's Exact Test (one-sided) $p < 0.001$) for Either Institution ACT mean/median ≥ 25.5 or Institution SAT mean/median 75th pctl ≥ 1195 cut-point met the criterion for inclusion in a multi-variable analysis of potential predictors.

Table D.4 Either Institution ACT 75th pctl ≥ 26 or Institution SAT mean/median ≥ 1132.5 for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
Either ACT 75th pctl ≥ 26 or SAT mean/median ≥ 1132.5	52	9
Neither ACT 75th pctl ≥ 26 nor SAT mean/median ≥ 1132.5	41	32
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.56 (95% CI: 0.468 – 0.66)	Sp = 0.78 (95% CI: 0.63 – 0.88)	
OR = 4.51 (95% CI: 1.94 – 10.50)	RFS = 1.52 (95% CI: 1.24 – 1.86)	

The OR of 4.51 (Fisher's Exact Test (one-sided) $p < 0.001$) for Either Institution ACT 75th pctl ≥ 26 or Institution SAT mean/median ≥ 1132.5 cut-point met the criterion for inclusion in a multi-variable analysis of potential predictors.

Table D.5 Either Institution ACT 80th pctl ≥ 27 or Institution SAT mean/median ≥ 1132.5 for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
Either ACT 80th pctl ≥ 27 or SAT mean/median ≥ 1132.5	52	9
Neither ACT 80th pctl ≥ 27 nor SAT mean/median ≥ 1132.5	42	33
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.55 (95% CI: 0.45 – 0.65)	Sp = 0.79 (95% CI: 0.64 – 0.88)	
OR = 4.54 (95% CI: 1.96 – 10.53)	RFS = 1.52 (95% CI: 1.24 – 1.87)	

The OR of 4.54 (Fisher's Exact Test (one-sided) $p < 0.001$) for Either Institution ACT 80th pctl ≥ 27 or Institution SAT mean/median ≥ 1132.5 cut-point met the criterion for inclusion in a multi-variable analysis of potential predictors.

Table D.6 Either Institution ACT 75th pctl ≥ 26 or Institution SAT 75th pctl ≥ 1195 for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
Either ACT 75th pctl ≥ 26 or SAT 75th pctl ≥ 1195	44	5
Neither ACT 75th pctl ≥ 26 nor SAT 75th pctl ≥ 1195	49	36
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.47 (95% CI: 0.38 – 0.57)	Sp = 0.88 (95% CI: 0.75 – 0.95)	
OR = 6.47 (95% CI: 2.33 – 17.93)	RFS = 1.56 (95% CI: 1.27 – 1.91)	

The OR of 6.47 (Fisher's Exact Test (one-sided) $p < 0.001$) for Either Institution ACT 75th pctl ≥ 26 or Institution SAT 75th pctl ≥ 1195 cut-point met the criterion for inclusion in a multi-variable analysis of potential predictors.

Table D.7 Either Institution ACT 80th pctl ≥ 27 or Institution SAT 80th pctl ≥ 1238 for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
Either ACT 80th pctl ≥ 27 or SAT 80th pctl ≥ 1238	30	3
Neither ACT 80th pctl ≥ 27 nor SAT 80th pctl ≥ 1238	64	39
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.32 (95% CI: 0.23 – 0.42)	Sp = 0.93 (95% CI: 0.81 – 0.98)	
OR = 6.09 (95% CI: 1.74 – 21.31)	RFS = 1.46 (95% CI: 1.19 – 1.79)	

The OR of 6.09 (Fisher's Exact Test (one-sided) $p < 0.001$) for Either Institution ACT 80th pctl ≥ 27 or Institution SAT 80th pctl ≥ 1238 cut-point met the criterion for inclusion in a multi-variable analysis of potential predictors.

Table D.8 Either Institution ACT mean/median ≥ 25.5 or an Institution SAT mean/median 80^{th} pctl ≥ 1238 for prediction of first-year gGPA (≥ 3.45)

	1^{st} Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
Either ACT 80^{th} pctl ≥ 27 or SAT 75^{th} pctl ≥ 1238	31	3
Neither ACT 80^{th} pctl ≥ 27 nor SAT 75^{th} pctl ≥ 1238	63	39
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.33 (95% CI: 0.24 – 0.43)	Sp = 0.93 (95% CI: 0.81 – 0.98)	
OR = 6.40 (95% CI: 1.83 – 22.34)	RFS = 1.48 (95% CI: 1.20 – 1.81)	

The OR of 6.40 (Fisher's Exact Test (one-sided) $p < 0.001$) for Either Institution ACT mean/median ≥ 25.5 or an Institution SAT mean/median 80^{th} pctl ≥ 1238 cut-point met the criterion for inclusion in a multi-variable analysis of potential predictors.

APPENDIX E

UNIVARIABLE ANALYSES AND 2 x 2 CROSS TABULATION TABLES FOR POTENTIAL PREDICTORS RELATED TO FIRST-YEAR SUCCESS (≥ 3.45)

APPENDIX E

Univariable analysis results for each of the potential predictors of first-year success (gGPA (≥ 3.45)) are provided in Figures E.1 to E.8 and Tables E.1 to E.11.

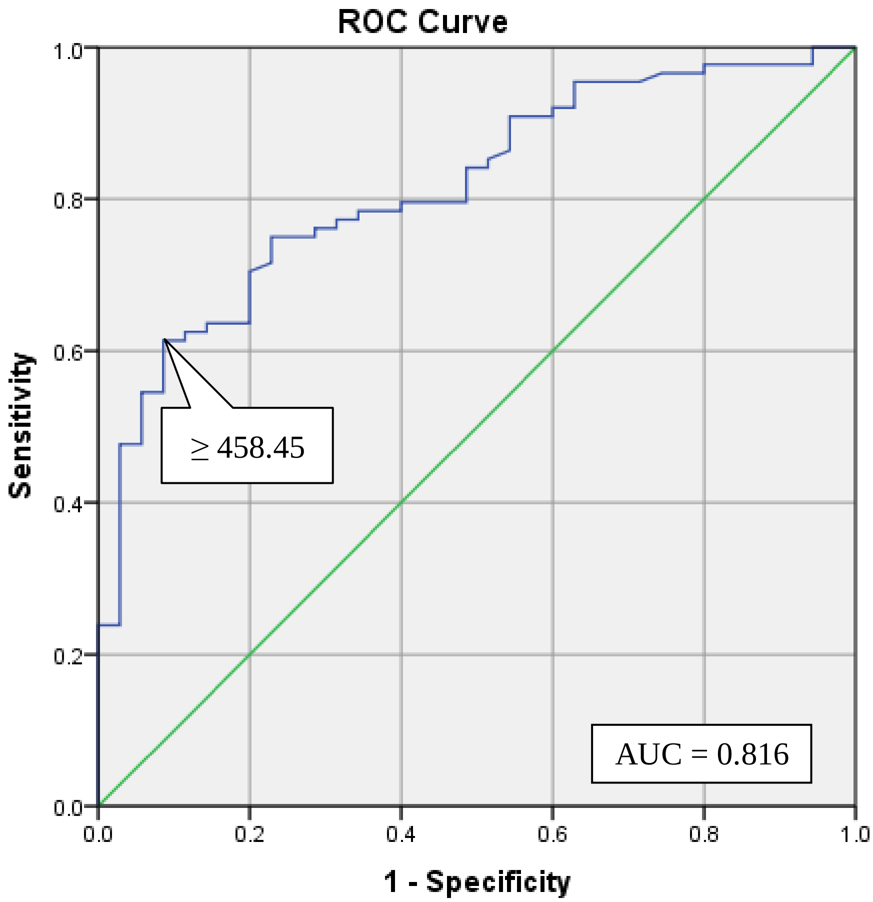

Figure E.1 ROC curve with identification of the optimum cut-point for Biderman's Formula Score for prediction of first-year gGPA (≥ 3.45)

Table E.1 Biderman's Formula Score for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
≥ 458.45	54	3
< 458.45	34	32
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.61 (95% CI: 0.51 – 0.71)	Sp = 0.91 (95% CI: 0.78 – 0.97)	
Youden's Index = 0.528		
OR = 16.94 (95% CI: 4.81 – 59.66)	RFS = 1.84 (95% CI: 1.50 – 2.25)	

A student in the GATP who had a Biderman's Formula Score of ≥ 458.45 had 16.94 times greater odds to be successful in the GATP than the odds for someone who had a Biderman's Formula Score of < 458.45.

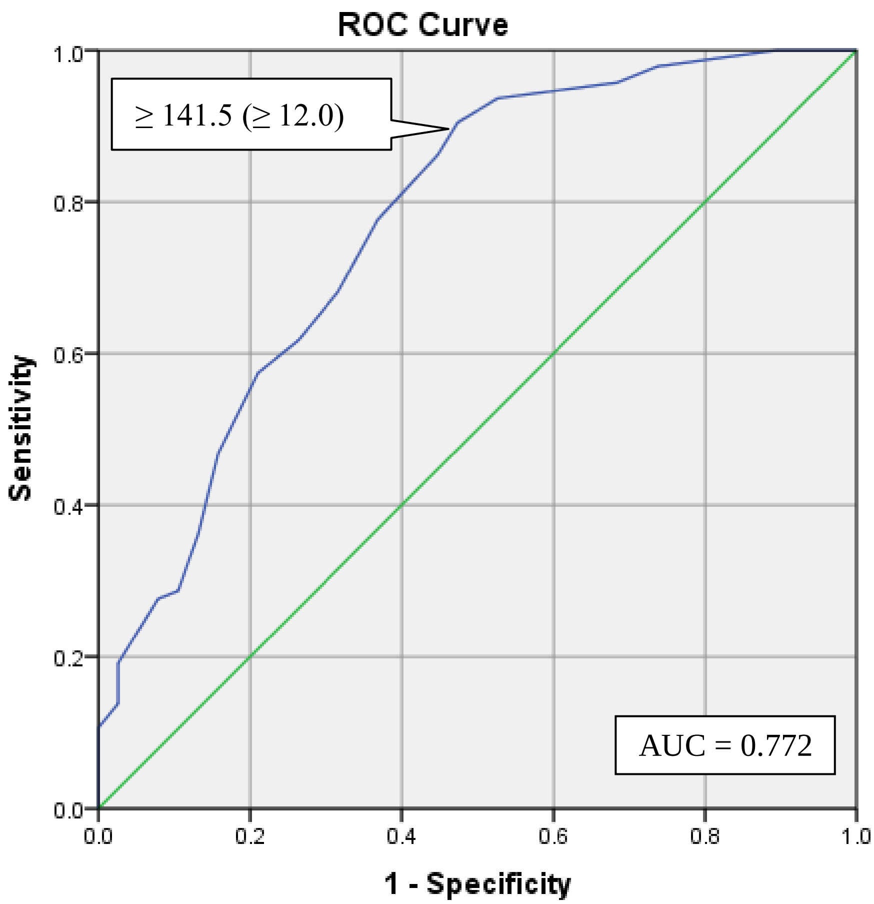

Figure E.2 ROC curve with identification of the optimum cut-point for GREq (PR) for prediction of first-year gGPA (≥ 3.45)

Table E.2 GREq (PR) for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
≥ 141.5 (12.0)	85	18
< 141.5 (12.0)	9	20
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.90 (95% CI: 0.838 – 0.95)	Sp = 0.53 (95% CI: 0.37 – 0.68)	
Youden's Index = 0.430		
OR = 10.49 (95% CI: 4.11 – 26.78)	RFS = 2.66 (95% CI: 2.17 – 3.26)	

A student in the GATP who had a GREq score of ≥ 141.5 (PR ≥ 12.0), had 10.49 times greater odds to be successful in the GATP than the odds for someone who had a GREq score < 141.5 (PR < 12.0).

Table E.3 Taking calculus as an undergraduate for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
Took calculus	41	3
Did not take calculus	53	39
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.44 (95% CI: 0.34 – 0.54)	Sp = 0.93 (95% CI: 0.81 – 0.98)	
OR = 10.06 (95% CI: 2.90 – 34.86)	RFS = 1.62 (95% CI: 1.32 – 1.98)	

A student in the GATP who took calculus as an undergraduate had 10.06 times greater odds to be successful in the GATP than the odds of someone who did not take calculus.

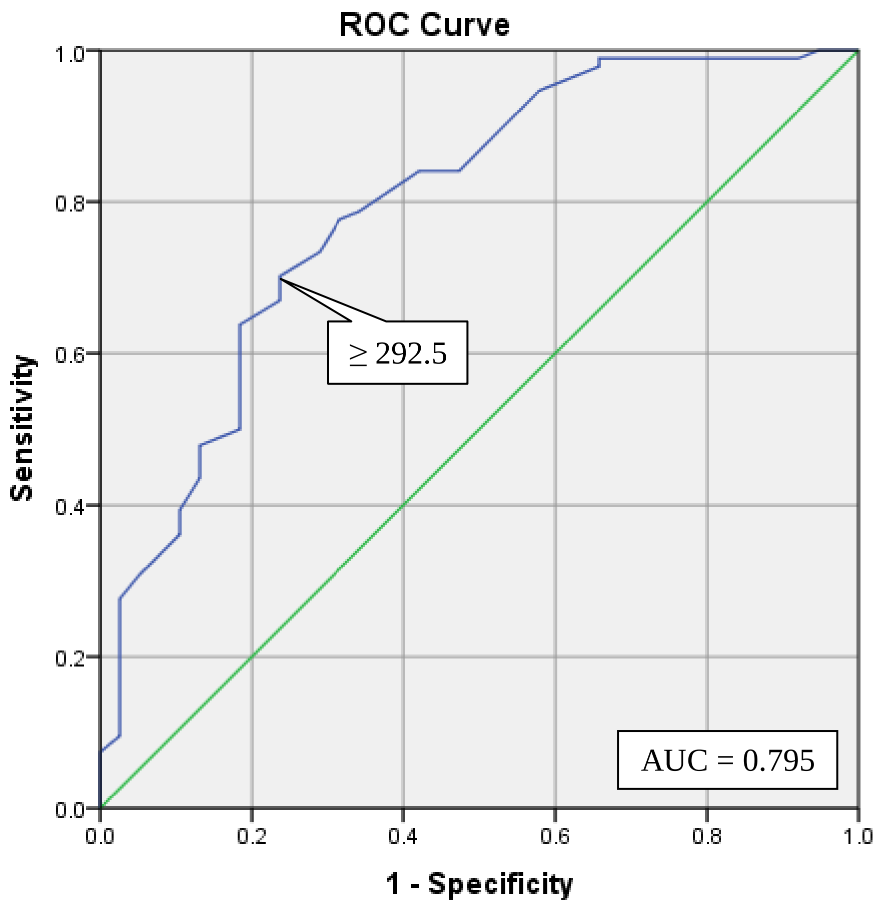

Figure E.3 ROC curve with identification of optimum cut-point for GRE Composite Score for prediction of first-year gGPA (≥ 3.45)

Table E.4 GRE Composite Score for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
≥ 292.5	66	9
< 292.5	28	29
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.70 (95% CI: 0.60 – 0.79)	Sp = 0.76 (95% CI: 0.61 – 0.87)	
Youden's Index = 0.465		
OR = 7.60 (95% CI: 3.19 – 10.11)	RFS = 1.79 (95% CI: 1.46 – 2.20)	

A student in the GATP who had a GRE Composite Score ≥ 292.5 had 7.60 times greater odds of being successful in the GATP than the odds for someone who had a GRE Composite Score < 292.5.

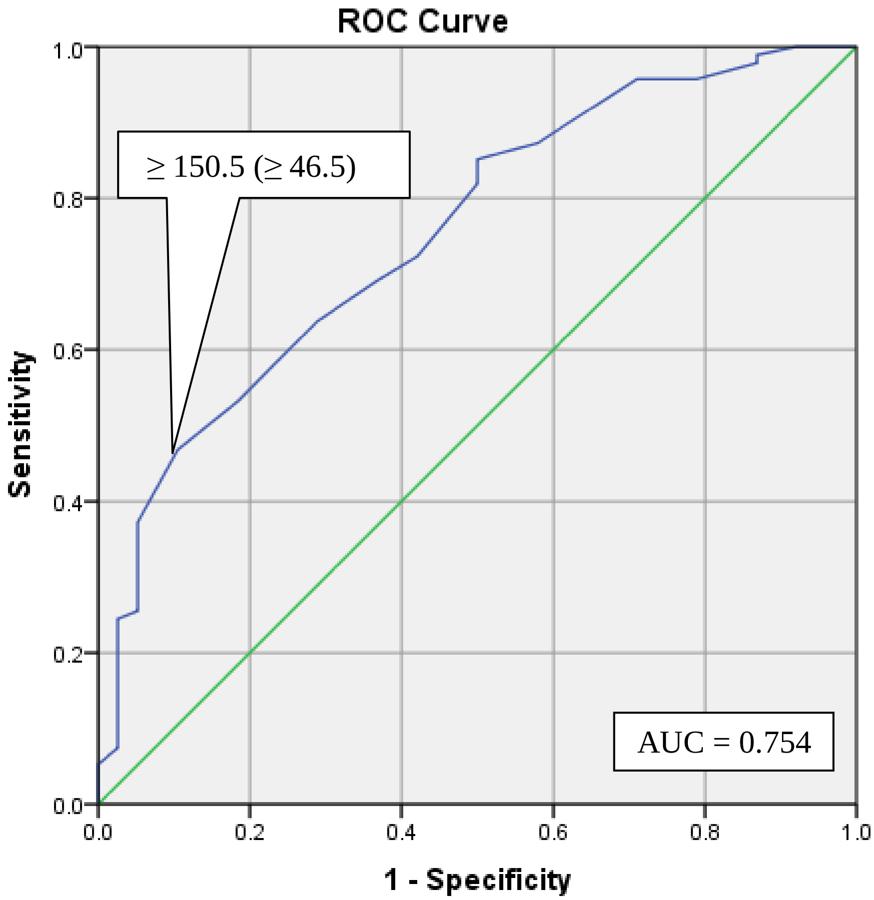

Figure E.4 ROC curve with identification of the optimum cut-point for GREv (PR) for prediction of first-year gGPA (≥ 3.45)

Table E.5 GREv (PR) for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
≥ 150.5 (46.5)	44	4
< 150.5 (46.5)	50	34
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.47 (95% CI: 0.37 – 0.57)	Sp = 0.89 (95% CI: 0.76 – 0.96)	
Youden's Index = 0.363		
OR = 7.48 (95% CI: 2.46 – 22.75)	RFS = 1.54 (95% CI: 1.26 – 1.89)	

A student in the GATP who had a GREv Score ≥ 150.5 (PR ≥46.5) had 7.48 times greater odds to be successful in the GATP than the odds for someone who had a GREv Score < 150.5 (PR < 46.5).

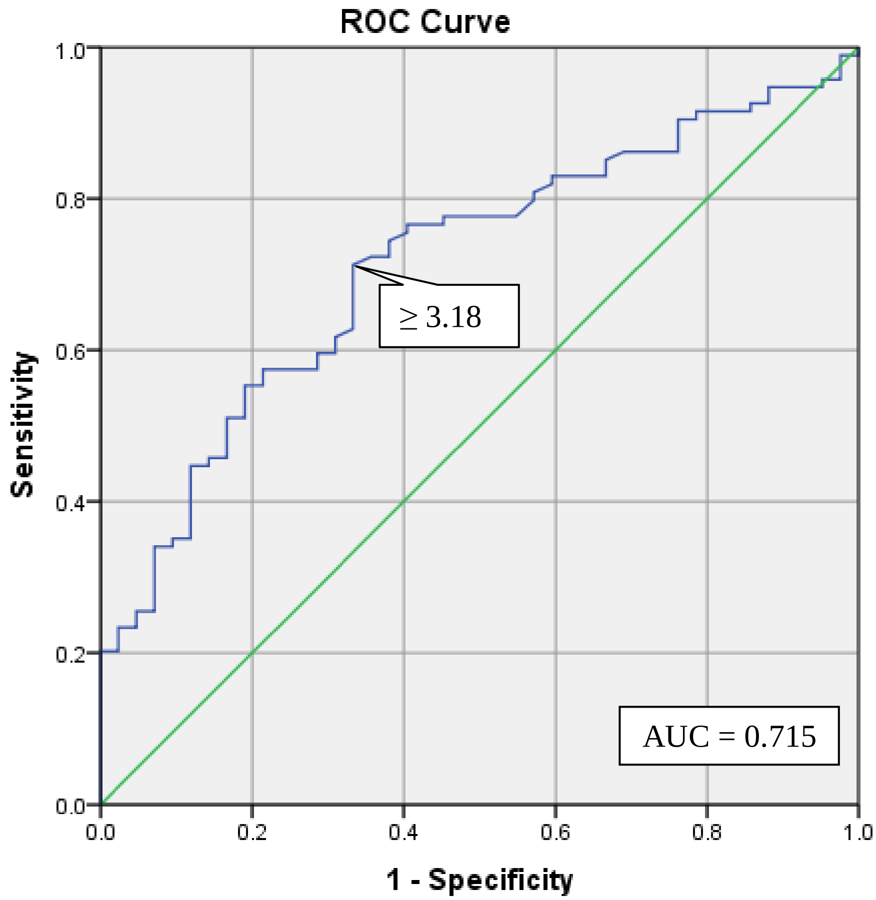

Figure E.5 ROC curve with identification of the optimum cut-point for uGPA for prediction of first-year gGPA (≥ 3.45)

Table E.6 uGPA for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
≥ 3.18	68	15
< 3.18	26	27
Fisher's Exact Test (one-sided) $p < 0.001$		
Sn = 0.72 (95% CI: 0.63 – 0.80)	Sp = 0.64 (95% CI: 0.49 – 0.77)	
Youden's Index = 0.380		
OR = 4.71 (95% CI: 2.17 – 10.23)	RFS = 1.67 (95% CI: 1.36 – 2.05)	

A student in the GATP who had an uGPA ≥ 3.18 had 4.71 times greater odds of being successful in the GATP than the odds for someone who had an uGPA < 3.18.

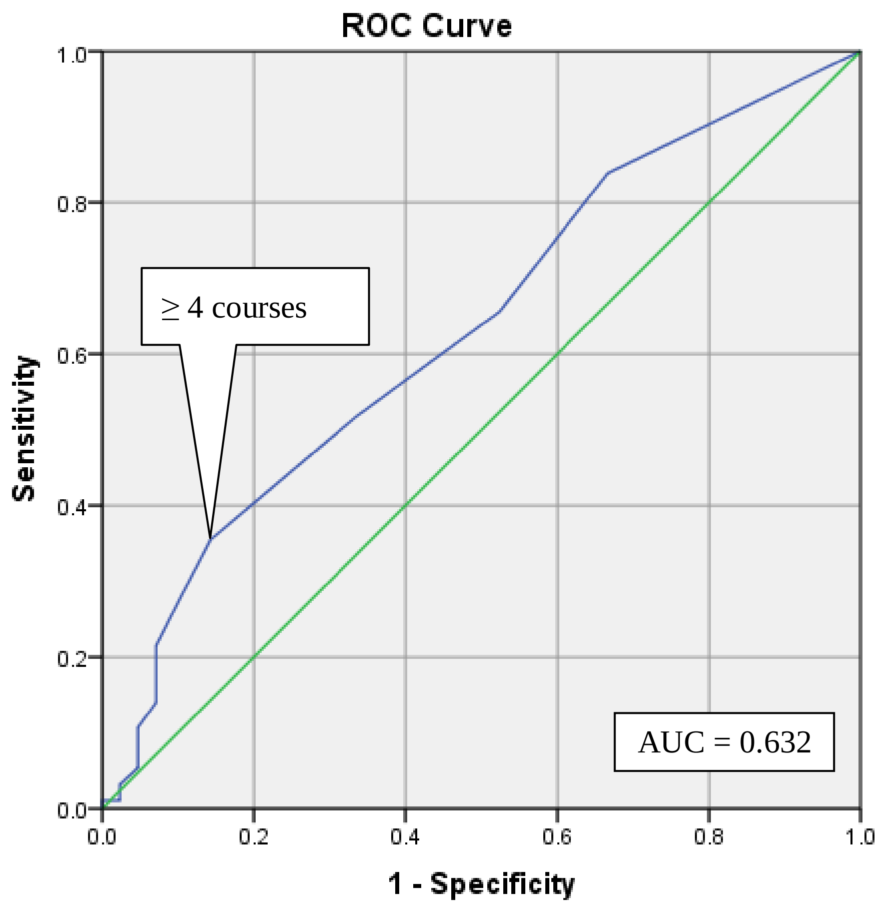

Figure E.6 ROC curve with identification of the optimum cut-point for the number of advanced math and science courses for prediction of first-year gGPA (≥ 3.45)

Table E.7 Number of advanced math and science courses for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
≥ 4 courses	33	6
< 4 courses	60	36
Fisher's Exact Test (one-sided) $p = 0.009$		
Sn = 0.36 (95% CI: 0.27 – 0.46)	Sp = 0.86 (95% CI: 0.72 – 0.93)	
Youden's Index = 0.212		
OR = 3.30 (95% CI: 1.26 – 8.65)	RFS = 1.35 (95% CI: 1.11 – 1.66)	

The OR of 3.30 (Fisher's Exact Test (one-sided) $p = 0.009$) for the number of advanced math and science courses cut-point of four courses met the criterion for inclusion in a multi-variable analysis of potential predictors.

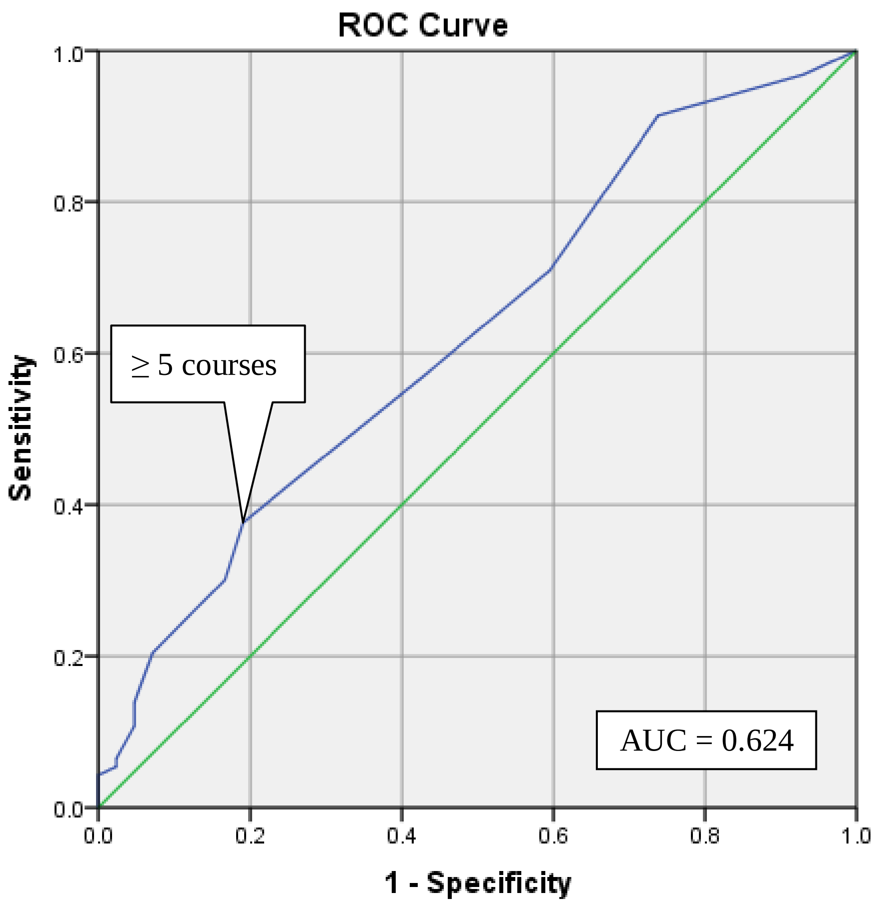

Figure E.7 ROC curve with identification of the optimum cut-point for the number of advanced courses for prediction of first-year gGPA (≥ 3.45)

Table E.8 Number of advanced courses for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
≥ 5 courses	35	8
< 5 courses	58	34
Fisher's Exact Test (one-sided) *p* = 0.045		
Sn = 0.38 (95% CI: 0.29 – 0.48)	Sp = 0.81 (95% CI: 0.67 – 0.90)	
Youden's Index = 0.186		
OR = 2.56 (95% CI: 1.07 – 6.17)	RFS = 1.29 (95% CI: 1.05 – 1.58)	

A student in the GATP who took five or more advanced science, math and athletic training courses during their undergraduate years had 2.56 times greater odds of being successful in the GATP than the odds for someone who took less than five advanced science, math and athletic training courses.

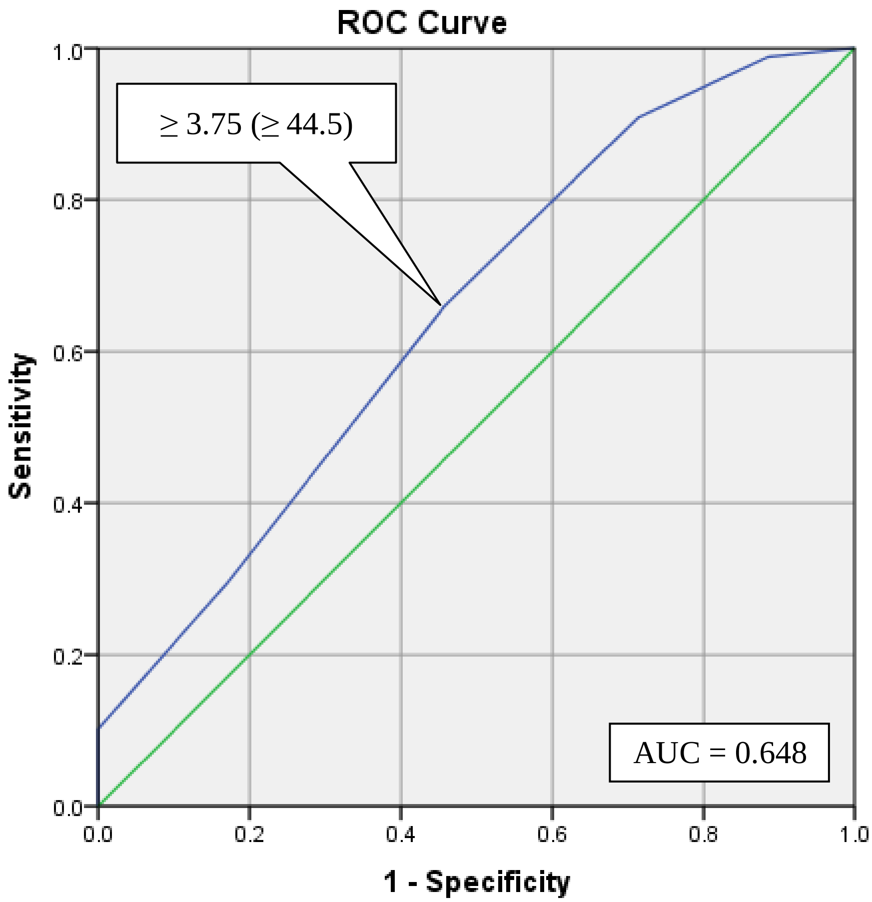

Figure E.8 ROC curve with identification of the optimum cut-point for GREwr (PR) for prediction of first-year gGPA (≥ 3.45)

Table E.9 GREwr (PR) for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
≥ 3.75 (44.5)	58	16
< 3.75 (44.5)	30	19
Fisher's Exact Test (one-sided) $p = 0.044$		
Sn = 0.66 (95% CI: 0.56 – 0.750)	Sp = 0.54 (95% CI: 0.38 – 0.70)	
Youden's Index = 0.202		
OR = 2.30 (95% CI: 1.03 – 5.01)	RFS = 1.28 (95% CI: 1.04 – 1.57)	

A student in the GATP who had a GREwr Score ≥ 3.75 (PR ≥ 44.5) had 2.30 times greater odds to be successful in the GATP than the odds for someone who had a GREwr score < 3.75 (PR < 44.5.)

Table E.10 Student graduated from a Research Intensive Institution for prediction of first-year gGPA (≥ 3.45)

	1st Year gGPA ≥ 3.45	**1st Year gGPA < 3.45**
Graduated from a Research Intensive Institution	43	14
Did not graduate from a Research Intensive Institution	51	28
Fisher's Exact Test (one-sided) $p = 0.121$		
Sn = 0.46 (95% CI: 0.36 – 0.56)	Sp = 0.67 (95% CI: 0.51 – 0.79)	
OR = 1.69 (95% CI: 0.79 – 3.60)	RFS = 1.17 (95% CI: 0.95 – 1.43)	

A student in the GATP who graduate from a research intensive institution had 1.69 times greater odds of being successful in the GATP than the odds for someone who did not graduate from a research intensive institution.

Table E.11 Taking physics as an undergraduate for prediction of first-year gGPA (≥ 3.45)

<table>
<tr><th></th><th colspan="2">1st Year gGPA ≥ 3.45</th><th>1st Year gGPA < 3.45</th></tr>
<tr><td>Took physics</td><td colspan="2">54</td><td>20</td></tr>
<tr><td>Did not take physics</td><td colspan="2">39</td><td>22</td></tr>
<tr><td colspan="4">Fisher's Exact Test (one-sided) p = 0.173</td></tr>
<tr><td colspan="2">Sn = 0.58 (95% CI: 0.48 – 0.68)</td><td colspan="2">Sp = 0.52 (95% CI: 0.38 – 0.67)</td></tr>
<tr><td colspan="2">OR = 1.52 (95% CI: 0.73 – 3.17)</td><td colspan="2">RFS = 1.14 (95% CI: 0.93 – 1.40)</td></tr>
</table>

A student in the GATP who had taken physics as an undergraduate had 1.52 times greater odds to be successful in the GATP than the odds for someone who did not take physics.

VITA

Scott Bruce was born in raised in Western Pennsylvania. Upon graduating high school he attended East Stroudsburg University in the Pocono Mountains of PA. He got a degree in Health, Physical Education, Recreation and Dance with a concentration in Athletic Training. Scott's first job was as the Head Athletic Trainer at Geneva College in Beaver Falls, PA, only a few miles from where he grew up in Freedom, PA. After working there for three years he returned to graduate school at Eastern Illinois University where he earned a Master's degree in Physical Education with a concentration in Sports Medicine.

After getting his degree from Eastern Illinois Scott married the former Jana Foust. They then moved to Adrian, MI and Scott's new job at Adrian College as their Head Athletic Trainer. He remained in that position for four and a half years until a "dream job" of sorts opened up in 1992. He was named the new Head Athletic Trainer at Slippery Rock University and he and Jana move home once again. While at Slippery Rock, Jana and Scott had two beautiful children, Patrick in 1994 and Allison in 1996.

The real "dream job" happened in 2001 when Scott got to live his dream of being an athletic trainer for a Division I football team. He became the assistant athletic trainer for football at the University of Miami. In his first-year at "The U" they won a Bowl Championship Series National Championship when the team went 12-0 and beat the University of Nebraska in the Rose Bowl. The following year, the team should have won a second national title, but did not.

Scott stayed one more year before leaving to come home and be the Head Athletic Trainer at California University of PA. Although the job brought Scott back home, (90 miles from his parents), he realized that the fit was not right and moved onto the University of Tennessee at Chattanooga for a full-time faculty position. Scott completed his Ed.D. degree in Learning and Leadership in 2014. Scott loves to garden, read, and play golf in his free time.